



Caractérisation acoustique des relations entre les mouvements biologiques et la perception sonore : application au contrôle de la synthèse et à l'apprentissage de gestes

Etienne Thoret

► To cite this version:

Etienne Thoret. Caractérisation acoustique des relations entre les mouvements biologiques et la perception sonore : application au contrôle de la synthèse et à l'apprentissage de gestes. Acoustique [physics.class-ph]. Aix-Marseille Université, 2014. Français. NNT : . tel-01105122

HAL Id: tel-01105122

<https://theses.hal.science/tel-01105122>

Submitted on 19 Jan 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Aix-Marseille Université

ED353 - Sciences pour l'ingénieur : Mécanique, Physique, Micro et Nanoélectronique

Laboratoire de Mécanique et d'Acoustique - CNRS - UPR 7051

Thèse présentée pour obtenir le grade universitaire de docteur

Discipline : Acoustique

Etienne THORET

Caractérisation acoustique des relations entre les mouvements biologiques et
la perception sonore : application au contrôle de la synthèse et à
l'apprentissage de gestes

Soutenue le 19 décembre 2014 devant le jury composé de :

Christophe Bourdin	Aix-Marseille Université - CNRS - ISM - UMR 7287	Président
Cathy Craig	Queen's University - Belfast	Rapporteur
Stephen E. McAdams	McGill University - Montreal	Rapporteur
Edouard Gentaz	Université de Genève	Examineur
Lionel Bringoux	Aix-Marseille Université - CNRS - ISM - UMR 7287	Invité
Richard Kronland-Martinet	CNRS - LMA - UPR 7051 - Aix-Marseille Université	Invité
Jean-Luc Velay	CNRS - LNC - UMR 7291 - Aix-Marseille Université	Invité
Sølvi Ystad	CNRS - LMA - UPR 7051 - Aix-Marseille Université	Directrice
Mitsuko Aramaki	CNRS - LMA - UPR 7051 - Aix-Marseille Université	Co-directrice

Remerciements

Un peu comme l'artisanat, la recherche est une forme de compagnonnage et ne peut perdurer sans la transmission du savoir des aînés. Je tiens ainsi à remercier en premier lieu mes trois encadrants, Sølvi (Ystad), Mitsuko (Aramaki), et Richard (Kronland-Martinet) pour leurs qualités, aussi bien scientifiques qu'humaines, qui m'ont permis de réaliser cette thèse dans un cadre scientifique et avec des moyens dont rêveraient tout étudiant en thèse. Je vous remercie pour le temps passé à lire, relire et corriger articles, chapitres... mais aussi pour votre ouverture d'esprit et la liberté que vous m'avez accordée tout au long de ces trois années. Je souhaite à tout étudiant de pouvoir réaliser une thèse dans ces conditions.

Je tiens également à remercier toutes les personnes avec qui j'ai pu travailler. En particulier, merci Lionel (Bringoux) et Jean-Luc (Velay) de m'avoir aidé et aiguillé dans le monde du mouvement et de la perception qui m'était complètement inconnu il y a encore trois ans, cette thèse n'aurait pas pu être menée à bien sans vous. Merci aussi à Jérémy (Danna) pour les échanges toujours agréables.

Comme dans l'artisanat, la critique est nécessaire et je remercie ainsi sincèrement tous les membres de mon jury d'avoir lu, commenté, et critiqué ces travaux de façon toujours constructive.

Ce paragraphe est un peu court pour remercier Simon (Conan), mon jumeau de thèse, pour le soutien mutuel, les échanges, et les bons moments partagés ces quatre dernières années. J'espère que l'on continuera encore longtemps à travailler ensemble !

Un (très) grand merci à mon coloc de bureau et ami Charles Max/MSP (Gondre) pour ton aide technique permanente et m'avoir fait comprendre que convoluer et filtrer n'étaient parfois pas si différents.

Merci aussi Soizic (Terrien) pour ton enthousiasme permanent, mais surtout pour m'avoir fait rejouer de la flute à bec bien que tu aies démontré que je suis toujours un non musicien non instrumentiste professionnel moins bon qu'une bouche artificielle !

Merci à Sebastien (Denjean) aka *l'induss*, qui a tant donné bien avant l'envie, pour tous les bons moments passés à Marseille en 2013.

Je remercie particulièrement Adrien (Merer) pour avoir initié les études sur les relations entre son et mouvement dans l'équipe. Je suis aussi *extremely thankful to* Thierry (Voinier) pour m'avoir initié aux subtilités du franglais mais surtout aux subtilités du signal audio toujours avec patience et pédagogie. Un remerciement particulier au gars Clément (François) pour son paternalisme et surtout de m'avoir incité à mettre le nez dans les bouquins de Varela et dans le Diplomatico. Merci aussi aux différents permanents du bâtiment P (Sabine (Meunier), Sophie (Savel), Michelle (Laurent), Jacques (Chatron), Olivier (Macherey)) pour leur bonne humeur et les échanges toujours sympathiques.

Un clin d'oeil à Lennie (Gandemer), Charly (Verron), Adrien (Sirdey), Gauthier (Real), Pyo (Michaud), Quentin (Mesnildrey) et Harold (Omer) pour les bonnes tranches de rigolades. Un remerciement particulier à Thomas (Bordonné) pour m'avoir supporté comme bourreau de stage. Il y a beaucoup de passages dans les labos et je tiens à saluer tous les doctorants, post-docs, et stagiaires qui sont passés pas loin du Bâtiment AA depuis trois ans, un salut plus particulier à Anaik (Olivero), Gaëtan (Parsehian), Fabrice (Silva), Jean-Baptiste (Doc), et Sami (Karkar). Un autre clin d'oeil à Paul (Luizard) et Romain *Maro* (Dumoulin) avec qui j'ai fait mes premiers pas dans le monde de l'acoustique.

Toujours comme dans l'artisanat, une fois le travail terminé il faut bien s'occuper. Je ne m'hasarderai pas à remercier toutes les personnes rencontrées à Marseille, je glisse juste un petit clin d'oeil à Saïd du Bar de la Plaine. Un clin d'oeil particulier à la famille Moulin, aux hérons cendrés, et aux poulpes. Merci aussi à mes potes d'Orléans pour l'accueil lors des pit-stops parisiens, marocains, et niçois, mais aussi aux squatteurs de l'été, à Marie et Flo, et Jérôme (Nika).

Enfin je remercie mon frère, mes parents, mes grands-parents, les Wittner, et ma famille pour leur affection et leur soutien inconditionnel (Un remerciement particulier à mon instit' de Maman pour la relecture intégrale du manuscrit !).

*Pour leurs inspirations diverses,
je dédie ces travaux à mes deux grands-pères*

Table des matières

Introduction	10
1 Etat de l'art & Problématiques	15
1.1 Les différentes approches dans la psychologie de la perception et de l'action	16
1.1.1 L'approche cognitive	17
1.1.2 <i>Au commencement était l'action</i> : les approches sensori-motrices de la perception	18
1.1.3 La perception : <i>une action simulée</i>	21
1.2 Approche écologique de la perception auditive	23
1.2.1 Perception auditive des objets : les <i>invariants structuraux</i>	24
1.2.2 Perception auditive des actions : les <i>invariants transformationnels</i>	25
1.2.3 Perception auditive des mouvements	25
1.3 Production motrice des mouvements biologiques	27
1.3.1 Invariants cinématiques des mouvements biologiques	28
1.3.2 Dynamique et coordination des mouvements biologiques	29
1.4 Perception des mouvements biologiques	31
1.4.1 La loi en puissance 1/3 contraint la perception visuelle de la vitesse.	32
1.4.2 La loi en puissance 1/3 contraint la perception de la géométrie d'une trajectoire.	34
1.5 Perception et catégorisation des formes géométriques	36
1.5.1 La perception d'une forme géométrique : une action simulée	36
1.5.2 Catégorisation des formes géométriques : une expérience visuelle ?	37
1.6 Contexte d'étude : analyse par la synthèse de sons de frottement	38
1.6.1 Acoustique du frottement	39
1.6.2 Synthèse et contrôle perceptif des sons de frottement	41
1.7 Enjeux	46
1.7.1 Enjeux théoriques	47
1.7.2 Enjeux applicatifs	48
2 Du son aux formes géométriques : perception auditive des mouvements graphiques	49
2.1 Introduction	49
2.2 Synthèse sons de frottement	51
2.3 Expérience 1 - Reconnaissance auditive des mouvements biologiques	51
2.3.1 Méthode	52
2.3.2 Résultats	53
2.3.3 Discussion	53
2.4 Expérience 2 - Discrimination auditive de formes géométriques distinctes	54
2.4.1 Méthode	55

2.4.2	Résultats	59
2.4.3	Discussion	59
2.5	Expérience 3 - Discrimination auditive de formes géométriques proches	61
2.5.1	Méthode	61
2.5.2	Résultats	62
2.5.3	Discussion	63
2.6	Le profil de vitesse : un invariant perceptif	64
2.6.1	Classification objective des formes en fonction des profils de vitesse	64
2.6.2	Comparaison entre classifications perceptives et mathématiques	65
2.6.3	Discussion	67
2.7	Discussion générale	67
3	Ecouter les sons produits par des mouvements biologiques révèle les prototypes géométriques	70
3.1	Introduction	71
3.2	Expérience 4 - Prototype visuel statique	76
3.2.1	Méthodes	76
3.2.2	Résultats	78
3.2.3	Discussion intermédiaire	78
3.3	Expérience 5 - Prototype moteur	80
3.3.1	Méthodes	80
3.3.2	Résultats	82
3.3.3	Discussion intermédiaire	82
3.4	Expérience 6 - Prototype visuel dynamique	83
3.4.1	Méthodes	84
3.4.2	Résultats	85
3.4.3	Discussion intermédiaire	86
3.5	Expérience 7 - Prototype sonore	87
3.5.1	Méthodes	87
3.5.2	Résultats	88
3.5.3	Discussion intermédiaire	89
3.6	Comparaison entre les différents prototypes	90
3.6.1	Excentricité	91
3.6.2	Variabilité des prototypes entre les modalités	91
3.6.3	Conclusion intermédiaire	92
3.7	Expérience 8 - Comparaison entre les métriques sonore et visuelle	93
3.7.1	Méthodes	93
3.7.2	Résultats	94
3.7.3	Discussion intermédiaire	96
3.8	Discussion générale	98


4	Voir des cercles, dessiner des ellipses : quand l'audition trompe la perception visuelle des mouvements biologiques	104
4.1	Introduction	104
4.2	Méthodes générales	107
4.2.1	Participants	107
4.2.2	Stimuli	107
4.2.3	Appareillage	110
4.2.4	Tâche	112
4.2.5	Analyse des données	112
4.2.6	Statistiques	113
4.3	Expérience préliminaire – Couplage audio-moteur	113
4.4	Expérience 9 – Couplage visuo-moteur	114
4.4.1	Résultats	115
4.4.2	Discussion	116
4.5	Expérience 10 – Couplage audio-visuo-moteur	117
4.5.1	Résultats	117
4.5.2	Discussion	120
4.6	Discussion générale	121
5	Vers la création de métaphores sonores et la sonification du geste d'écriture	125
5.1	Création de métaphores sonores évoquant un geste humain	126
5.2	Vers l'apprentissage d'un geste expert avec le son : une nouvelle écriture	128
5.2.1	Variables caractéristiques du geste d'écriture	128
5.2.2	Stratégie de sonification : la métaphore du couinement	128
5.2.3	Une preuve de concept : apprentissage d'un geste par la main non-dominante	133
5.3	Diagnostic et réhabilitation de la dysgraphie	136
5.3.1	La dysgraphie : une pathologie graphomotrice	137
5.3.2	Stratégie de sonification	137
5.3.3	Premières pistes pour la réhabilitation de la dysgraphie par le biais d'un retour sonore	137
5.4	Conclusion	137
6	Conclusion générale & Perspectives	139
6.1	Vers d'autres relations entre le son et le mouvement biologique	140
6.1.1	D'autres invariants associés aux gestes humains ?	140
6.1.2	D'autres invariants associés aux mouvements humains ?	141
6.1.3	Les émotions transmises par la musique : une simulation des mouvements biologiques ?	142
6.2	Vers des stratégies de sonification pour le guidage et l'apprentissage de gestes	143
6.2.1	Vers le guidage du geste par le son	144
6.2.2	Sonifier d'autres invariants ?	145
6.3	Vers un véritable paradigme d'étude des relations entre gestes et sons	145
6.3.1	Une nouvelle illusion audiovisuelle induite par l'audition	145
6.3.2	La perception auditive de la dynamique : une inférence bayésienne ?	146
6.3.3	Déterminer la contribution des interactions sensori-motrices dans l'émergence des prototypes géométriques	147

Postface	147
7 Annexes	I
7.1 Annexe 1 – Le mouvement de Lissajous respecte la loi en puissance $1/3$	I
7.2 Annexe 2 – Effet du mapping sur la synchronisation audio-motrice	III
7.3 Annexe 3 – Résultats détaillés du chapitre 3	XI
7.4 Annexe 4 – Effet du mapping sur la représentation mentale de la taille	XIII
7.5 Annexe 5 – Autres exemples de friction non linéaire	XVII
7.6 Annexe 6 – Publications et communications associées à la thèse	XIX
Bibliographie	XX

Exemples sonores et vidéos en ligne

Des exemples sonores, vidéos, et les stimuli utilisés dans toutes les expériences qui seront présentées dans cette thèse sont disponibles en ligne à l'adresse suivante :

<http://www.lma.cnrs-mrs.fr/%7Ekronland/TheseEThoret/>

Au cours du manuscrit, chaque exemple multimédia ou stimuli d'expérience accessible en ligne sera signalé par le symbole .

Introduction

Qui n'a jamais observé un bébé bouger spontanément en rythme avec le tempo d'un morceau de musique ? Des études anciennes ont fait état de ces effets dynamogéniques de la musique, et plus généralement des sons, sur les mouvements du corps (Fraisse *et al.*, 1949, 1953). De même, l'harmonie entre les mouvements du corps d'une danseuse et les variations temporelles d'un morceau de musique montre à quel point les mouvements du corps peuvent décrire précisément, et même transmettre des émotions, liées aux propriétés acoustiques des sons. Enfin, les mouvements d'un chef d'orchestre constituent un moyen efficace pour traduire les intentions musicales d'un compositeur à un groupe de musiciens afin qu'ils puissent jouer en harmonie. A l'inverse, les sons que nous gérons au quotidien, qu'ils soient musicaux ou non, impliquent nécessairement un mouvement du corps. Que nous frappions sur une barre de xylophone pour jouer de la musique ou sur une porte pour annoncer notre arrivée chez un ami, c'est bien le mouvement de notre main qui fait entrer en résonance la barre de métal du xylophone ou la porte. Il semble donc qu'il existe une relation intrinsèque entre nos gestes et les qualités spectro-temporelles des sons, ayant des conséquences sur nos perceptions, nos émotions, et nos actions. Comprendre ces relations est un enjeu aussi bien du point de vue fondamental, afin de mieux connaître les processus cognitifs ainsi que les caractéristiques acoustiques sous-jacents à la perception des sons, qu'appliqué, cela permettant d'envisager des applications diverses et variées utilisant les sons pour informer et guider nos comportements. Ces deux aspects constituent les objectifs principaux de cette thèse.

Du point de vue fondamental, la relation entre les gestes et la cognition a fait l'objet de nombreuses études dans des contextes théoriques différents. Des conceptions historiques ont adopté un point de vue constructiviste et ont abordé le problème de la formation des percepts en terme de formation de représentation du monde à l'échelle du cerveau. Des approches alternatives plus récentes ont elles proposé d'intégrer le corps et la motricité au sein même de la formation des percepts. L'approche écologique de la perception postule par exemple que l'action est guidée par des informations invariantes extraites des flux perceptifs captées par nos sens. Une approche plus récente, appelée approche éactive, propose même que plus que d'être guidée par la perception, c'est notre corps même qui est à la source de nos perceptions, et que nous incarnons le monde par le biais de cette boucle sensori-motrice. Ainsi, en accord avec ces approches, nombreuses sont les études qui se sont intéressées aux relations entre les mouvements des musiciens et les sons produits par les instruments. Il a ainsi été mis en évidence un couplage très fort entre les compétences motrices et les sons produits par les mécanismes physiques des instruments. Sous le nom d'*approche incarnée de la cognition musicale*, Leman (2008) a proposé qu'au-delà d'un certain niveau d'expertise, les musiciens interprètent les sons que produisent leurs instruments directement comme des gestes potentiels qu'ils doivent réaliser pour jouer correctement de leur instrument. Cette approche fait écho à de nombreuses études sur la perception et la production de la parole. *La théorie motrice de la perception de la parole* suggère ainsi que nous percevons les sons produits par la parole et accédons au lexique sémantique associé en interprétant les indices acoustiques comme des gestes phonatoires potentiels (Liberman et Mattingly,

1985). Du point de vue cognitif, les mots et ces gestes ne seraient donc qu'une seule et même entité, activée tantôt lors de la production vocale d'un mot et tantôt lors de sa perception par le système auditif. De même, pour la perception des sons du quotidien, une approche basée sur l'approche écologique propose un cadre formel pour l'étude de la perception auditive de ces sons (Gaver, 1993b; McAdams, 1993). Cette approche suggère que le processus de reconnaissance d'un événement sonore s'appuie principalement sur deux invariants acoustiques pertinents du point de vue perceptif, d'une part un invariant dit structurel, caractérisant la structure de l'objet en vibration comme son matériau ou sa taille, et d'autre part un invariant dit transformationnel, caractérisant l'action qui est en train d'être réalisée sur l'objet, par exemple le fait que l'objet est en train de rebondir ou d'être frotté.

Dans le domaine des sciences du mouvement, les mouvements biologiques ont été intensément étudiés et il est maintenant bien connu que ces mouvements ont des propriétés cinématiques très particulières aussi bien au niveau de leur production motrice que de leur perception visuelle et haptique. Plus précisément, quand nous dessinons naturellement une forme géométrique sur une feuille de papier, nous ralentissons et accélérons alternativement dans les parties les plus et les moins courbées de la forme respectivement sans même nous en rendre compte. Ces covariations entre vitesse du geste graphique et géométrie de la forme sont formalisées par la *loi en puissance 1/3* qui stipule que la vitesse tangentielle v_t au cours du temps est inversement proportionnelle à la puissance 1/3 de la courbure locale $C(t)$: $v_t(t) = KC(t)^{-1/3}$, où K est une constante (Viviani et Terzuolo, 1982; Lacquaniti *et al.*, 1983). Ces covariations entre cinématique et courbure contraignent également la perception visuelle des mouvements. Une expérience de Viviani et Stucchi (1992) a montré que pour que la vitesse du mouvement d'un point lumineux soit perçue comme constante, il est nécessaire que le point accélère dans les parties plates et ralentisse dans les parties les plus courbées de la trajectoire en respectant la loi en puissance 1/3. Cette illusion montre donc que les propriétés cinématiques des mouvements biologiques sont également intégrées par nos processus perceptifs et cognitifs.

Pour étudier les relations entre les mouvements biologiques et la perception sonore, nous nous sommes intéressés au cas des gestes graphiques. A première vue, dans la mesure où ces mouvements génèrent une trace visuelle mais aucun son spécifique, ils semblent plus naturellement associés à la vision qu'à l'audition. La qualité d'une écriture est en effet principalement jugée sur la qualité de la trace et non sur la qualité des sons produits. Cependant, si on prête attention aux sons produits quand une personne dessine, on peut entendre très distinctement le son du frottement sur le papier, en particulier quand la surface est très rugueuse. Les changements de vitesse du stylo produisent ainsi des variations de timbre qui sont directement reliées au geste réalisé par le scripteur. Le fait que les gestes graphiques respectent les règles des mouvements biologiques mais qu'ils produisent également du son fait de ces mouvements un paradigme idéal pour l'étude des relations entre gestes et sons. Ils présentent notamment l'avantage d'induire une relation continue entre le geste et le son produit correspondant à la friction entre la mine du stylo et la feuille. En particulier, on peut se demander si l'on est capable de reconnaître les mouvements qui respectent la loi en puissance 1/3 juste par l'écoute de ces sons ? De plus, on peut se demander si l'on peut reconnaître la forme qui a été dessinée juste à partir du son ? Au premier abord, ces questions peuvent paraître fortuites voire futiles, cependant, des réponses positives mettraient en évidence la relation extrêmement forte entre les variations de timbre d'un son et les propriétés cinématiques de nos gestes dans une situation où l'audition ne joue *a priori* aucun rôle.

Un autre avantage des gestes graphiques est qu'ils sont naturellement associés à une forme vi-

suelle, i.e. la forme dessinée. Ce contexte d'étude permet de confronter trois modalités sensorielles : la vision, la proprioception et l'audition. Ainsi, on peut explorer l'existence de forme géométrique prototypique, i.e. forme représentant le mieux une catégorie de forme du point de vue perceptif (Rosch, 1973), associées à une forme particulière par le biais de ces trois modalités et en particulier par le biais de la modalité auditive. L'utilisation de la modalité auditive pour étudier l'existence de forme prototypique est particulièrement originale et intéressante dans la mesure où elle permet d'aborder ce problème, relevant des sciences du mouvement et de la psychologie, sous l'angle de la perception des sons de frottement et de leurs propriétés acoustiques sous-jacentes. Cela permet en particulier d'évaluer l'importance de la composante dynamique du geste indépendamment de la géométrie des formes considérées ce qui n'est pas possible avec les approches déjà utilisées impliquant la vision ou la proprioception.

Enfin, la relation entre vision, audition, et proprioception peut être étudiée au regard des formes prototypiques mais également dans un contexte multimodal. Dans la mesure où les sons de frottement peuvent évoquer le mouvement de façon très forte, on peut se demander s'il n'est pas possible d'altérer, voire de modifier la perception et la reproduction motrice de la dynamique et de la géométrie d'un mouvement visuel en y ajoutant une dynamique sonore incongruente contrôlée. Comme cela a été évoqué précédemment, la perception et la reproduction de la vitesse et de la géométrie d'un mouvement visuel est naturellement biaisée par les règles des mouvements biologiques. En outre, si les variations de vitesse d'un mouvement sont efficacement transmises par les sons de frottement, ils semblent donc être des candidats sérieux pour altérer la perception et la reproduction motrice de la vitesse et de la géométrie d'un mouvement visuel.

Pour aborder ces problèmes, des enregistrements des sons de frottement produits par des scripteurs pourraient être utilisés. Cependant, il est difficile de demander à quelqu'un de dessiner une forme en respectant un profil de vitesse particulier. En effet, la variabilité inhérente à la production d'un mouvement rend difficile la génération de ce type de stimuli. De même que l'on peut contrôler le mouvement d'un point lumineux pour générer un stimulus visuel précis et calibré, la synthèse sonore numérique permet désormais de générer des sons de frottement réalistes à partir de paramètres cinématiques (e.g. profil de vitesse) et dynamiques (e.g. pression) spécifiques. Ainsi, l'utilisation d'un modèle de synthèse sonore de frottement offre la possibilité d'étudier les relations entre gestes et son du point de vue fondamental. En effet, les travaux réalisés au Laboratoire de Mécanique et d'Acoustique depuis une quinzaine d'années ont permis le développement de synthétiseurs s'appuyant sur l'approche écologique de la perception auditive des événements du quotidien et plus particulièrement sur le paradigme action → objet. Ce paradigme permet de modéliser et manipuler très intuitivement les ressentis perceptifs liés aux sons en fonction des invariants acoustiques et perceptifs reliés à l'action, i.e. transformationnels, et ceux reliés à l'objet, i.e. structuraux. On peut par exemple synthétiser très facilement un son de frottement en manipulant les caractéristiques cinématiques de l'objet en mouvement indépendamment de toutes les autres propriétés et surtout de contrôler parfaitement leurs conséquences acoustiques sur le son généré. Cet outil permet donc à la fois de contrôler précisément les propriétés des stimuli sonores en temps réel et d'envisager des protocoles interactifs où l'on peut manipuler les propriétés acoustiques d'un son en fonction de caractéristiques cinématiques comme la loi en puissance 1/3.

Dans cette thèse, nous aborderons donc le problème des relations entre gestes et sons en s'appuyant à la fois sur des connaissances issues des sciences du mouvement, et plus particulièrement des modèles cinématiques de gestes, et d'autre part sur des connaissances issues de la psychologie

cognitive, en particulier l'approche écologique de la perception auditive. L'utilisation de modèles de synthèse sonore est la pierre angulaire de ce travail et se trouve ainsi au milieu de ces problèmes spécifiques. Il est en effet essentiel de constater que la synthèse sonore et son contrôle sera à la fois outil d'exploration de la perception mais également notre problème de fond. En effet, si pour comprendre la façon dont nous percevons les gestes à travers les sons, les problématiques relevant de la psychologie et des sciences du mouvement peuvent trouver une utilité dans l'utilisation de synthétiseurs permettant de générer des stimuli calibrés, la création de ces synthétiseurs nécessitent eux la connaissance de la façon dont nous percevons les sons qu'ils produisent afin de pouvoir les contrôler de façon intuitive. Ainsi, cette thèse abordera des problématiques issues des sciences cognitives et des sciences du mouvement avec le regard de l'acousticien curieux, mais surtout nécessaires de comprendre comment les sons qu'il génère par synthèse sont perçus et interprétés. Cette problématique est ainsi pluridisciplinaire en soi et soulève des questions issues de nombreux domaines, en particulier de la psychologie cognitive de l'audition et des sciences du mouvement. Comme l'a proposé Jean-Claude Risset (2010), la pluridisciplinarité consiste un peu à *s'asseoir entre deux chaises, ce qui est particulièrement inconfortable en France*, et il convient donc d'aborder ces problématiques spécifiques à ces domaines de façon modeste en essayant de se conformer à leurs règles spécifiques aussi bien méthodologiques que terminologiques. Cependant, au-delà de ces difficultés, cette approche est très riche et permet également d'envisager de nombreuses applications inenvisageables si l'on se borne à un domaine spécifique. En effet, les éventuelles réponses que nous pourrions apporter aux questions fondamentales sur les relations entre sons et gestes permettront par exemple d'envisager le développement de nouveaux outils d'apprentissage de gestes et de réhabilitation de pathologies graphomotrices par le biais de la modalité auditive en utilisant des outils de synthèse sonore.

Un premier chapitre présentera ainsi une revue de littérature sur les connaissances issues de la psychologie et des sciences du mouvement nécessaires pour aborder le problème des relations entre mouvements biologiques et perception sonore. Cet état de l'art se conclura par la présentation d'un paradigme de synthèse sonore, aussi appelé paradigme action → objet, que nous utiliserons lors des différentes études et qui permet de contrôler précisément les paramètres acoustiques des sons de frottement en fonction d'un profil de vitesse spécifique et au regard des connaissances sur la façon dont nous percevons les sons. Dans le chapitre 2, une première étude évaluera dans quelle mesure il est possible de reconnaître les variations de vitesse caractéristiques des mouvements biologiques, i.e. la loi en puissance $1/3$, et de discriminer des formes géométriques sur la base des informations cinématiques transmises par les sons de frottement produits quand une personne dessine. Dans le chapitre 3, une seconde étude s'intéressera à l'existence de formes géométriques prototypiques notamment du point de vue de la modalité auditive, cela permettra d'évaluer l'importance des contraintes motrices dynamiques dans l'émergence de la géométrie des prototypes. Dans le chapitre 4, une troisième étude s'intéressera elle à évaluer dans quelle mesure la perception auditive peut altérer, voire modifier la perception visuelle et la reproduction motrice d'un mouvement visuel. Le chapitre 5 présentera lui des perspectives de création de sons inouïs basé sur le croisement entre invariants perceptifs, également appelés métaphores sonores, évoquant des situations a priori impossibles. Ce chapitre présentera également une perspective d'utilisation de la synthèse sonore comme moyen de guidage et de réhabilitation d'un geste graphique. Enfin, les conclusions et perspectives de ces études seront présentées dans un dernier chapitre.

Chapitre 1

Etat de l'art & Problématiques

Sommaire

1.1	Les différentes approches dans la psychologie de la perception et de l'action . . .	16
1.2	Approche écologique de la perception auditive	23
1.3	Production motrice des mouvements biologiques	27
1.4	Perception des mouvements biologiques	31
1.5	Perception et catégorisation des formes géométriques	36
1.6	Contexte d'étude : analyse par la synthèse de sons de frottement	38
1.7	Enjeux	46

Nous allons donc nous intéresser aux relations entre la perception sonore et les mouvements biologiques par le biais des sons de frottement produits quand une personne dessine. Afin d'aborder ce problème, il est nécessaire de comprendre comment l'information acoustique arrivant à l'oreille d'une personne est convertie puis analysée par le système auditif mais également de comprendre comment nous produisons et percevons les mouvements biologiques dans d'autres modalités que l'audition. Dans un premier temps, nous présenterons ainsi les rudiments des approches psychologiques abordant les problématiques de la perception et de l'action. Cette section s'appuiera essentiellement sur des exemples issus de la modalité visuelle, ce qui nous permettra de bien comprendre dans quel contexte nous allons aborder le problème des relations entre gestes et sons. Nous nous focaliserons dans un second temps sur la perception et la reconnaissance des événements sonores du quotidien au regard de l'approche écologique de la perception sonore. Une fois ce contexte théorique posé, nous présenterons les connaissances relevant de la perception visuelle et de la production des mouvements biologiques. Ainsi, nous nous rendrons compte que quasiment aucune étude n'a étudié la perception des mouvements biologiques du point de vue de l'audition. Enfin, nous présenterons le paradigme de synthèse sonore action → objet basé sur l'approche écologique. Ce paradigme de synthèse est au centre de toutes nos études et constitue l'originalité de tous les travaux qui seront présentés. Finalement, au regard des possibilités offertes par ce paradigme de synthèse et au regard des connaissances théoriques sur la perception et la production des mouvements biologiques, nous résumerons les enjeux et les différentes problématiques qui seront abordés dans les différents chapitres de cette thèse.

1.1 Les différentes approches dans la psychologie de la perception et de l'action

La façon dont les êtres humains perçoivent le monde sensible et dont ils interagissent avec est une question qui a été abordée depuis très longtemps. Les philosophes de l'antiquité comme Platon se questionnaient déjà sur la nature des *images* et des percepts transmis par nos sens puis interprétés par le cerveau. *L'allégorie de la caverne* est certainement un des premiers exemple d'un tel questionnement. Platon a imaginé une situation où des hommes prisonniers sont enfermés dans une caverne où ils n'ont jamais vu directement la lumière du jour. Ils ne peuvent découvrir le monde visuel que par les ombres produites par le peu de lumière qui parvient jusqu'à eux. De même ils ne découvrent le monde sonore que par les échos des sons qui arrivent à leurs oreilles (voir figure 1.1). Il fit alors l'hypothèse que si l'un des prisonniers était libéré et qu'il avait la possibilité de sortir de la caverne, il sera dans un premier temps incapable de percevoir le monde dans sa réalité sensible, en effet, sa construction du monde n'est alors basée que sur son expérience dans la caverne où les sensations qu'il a pu jusqu'à présent percevoir sont très différentes de celles auxquelles il est désormais confronté et qu'il ne sait interpréter. Platon imagine alors que le prisonnier pourra découvrir la nouvelle réalité sensible s'il réalise un effort de contemplation qui lui permettra de comprendre et interpréter les images et les sons qui arrivent à ses yeux et oreilles et qu'il n'a jusqu'alors jamais entendus. En postulant que le monde ne se découvre que par la contemplation, Platon distingue ainsi clairement le monde matériel de celui de l'esprit et émet l'hypothèse que la réalité sensible n'est qu'une illusion opposée au monde de l'esprit et des idées.

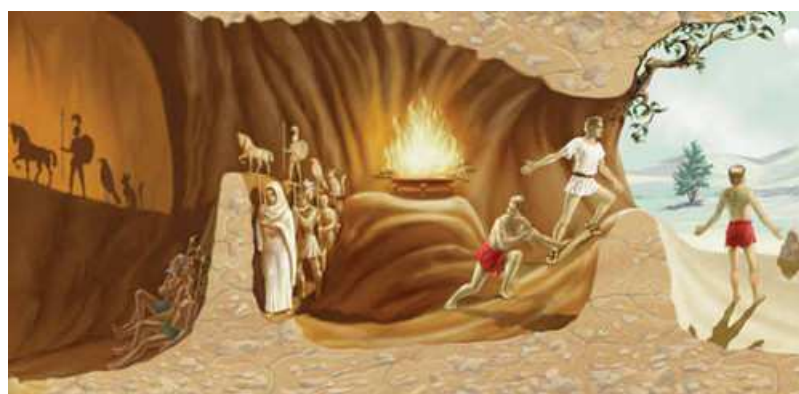


FIGURE 1.1 – Illustration représentant l'allégorie de la caverne : des prisonniers sont enfermés dans la grotte et ne peuvent découvrir le monde sensible que par les ombres qu'ils voient – Illustration issue de <http://www.john-uebersax.com/plato/myths/cave.htm>

Cette approche subjectiviste de la perception postule donc que le monde et l'esprit sont clairement séparés et que la *vérité* ne peut être atteinte que par l'introspection. Cette approche a été longtemps adoptée et il a fallu attendre le moyen-âge pour que le regard sur la façon dont l'esprit pouvait être relié au monde évolue (O'Regan, 2011). En s'intéressant au fonctionnement du système optique, Kepler et Descartes ont permis de faire considérablement évoluer les connaissances à ce sujet¹. Lors d'observations astronomiques à l'oeil nu, Kepler s'est rendu compte que les mesures de la taille de la lune ou du soleil variaient fortement en fonction de la façon dont les mesures étaient réalisées. Kepler (1604) fit donc l'hypothèse que les yeux n'étaient pas le siège de la vision et qu'ils étaient juste un *ins-*

1. Une revue précise de l'influence des idées des philosophes grecs sur la conception du fonctionnement du système visuel est présentée sur le site <http://nivea.psych.univ-paris5.fr/FeelingSupplements/AncientVisions.htm> en supplément de l'ouvrage *Why red doesn't sound like a bell?* de Kevin O'Regan (2011).

trument optique dont les données étaient ensuite analysées par le cerveau. Descartes (1972) confirma cette intuition de Kepler en constatant que l'image qui se forme sur la rétine est inversée par rapport au monde réel, voir figure 1.2. Cette observation confirma que l'oeil ne pouvait être le siège de la vision et qu'un traitement de l'image par le cerveau avait nécessairement lieu. Descartes proposa diverses explications physiologiques sur la façon dont le cerveau renversait ensuite l'image pour qu'elle corresponde à l'image de la réalité². Il fit en particulier l'hypothèse que *le siège de l'âme* se situait dans la glande pinéale située au milieu du cerveau juste au dessus du cervelet. Selon Descartes, le rôle de cette glande était d'interpréter les informations visuelles pour en donner une transcription cohérente au sujet. Cette vision de la perception, également appelée *homonculus* (*petit homme*) ou théâtre Cartésien chez les philosophes modernes, fait l'hypothèse qu'un agent intelligent *voit* ce qui est vu par le sujet et prend ensuite les décisions comportementales appropriées.

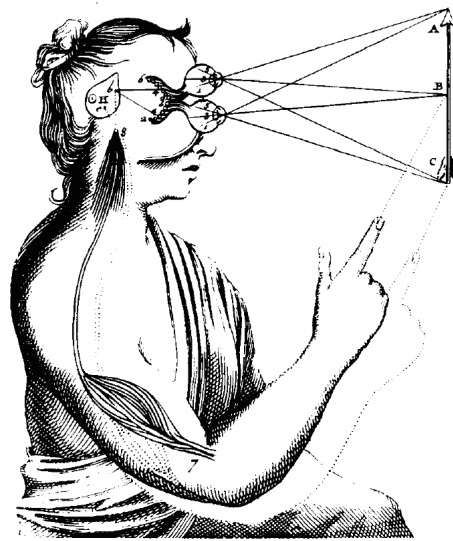


FIGURE 1.2 – *Le théâtre cartésien* - Modèle de la perception visuelle proposé par Descartes (1972) basé sur des processus afférents (*bottom-up*) et efférents (*top-down*) dont le fonctionnement est contrôlé au niveau de la glande pinéale - *homonculus*

1.1.1 L'approche cognitiviste

Les explications proposées par Descartes sont critiquables en particulier le postulat d'un agent intelligent qui n'est pas une solution au problème de la perception. En effet, supposer l'existence d'un tel agent pose alors la question de son fonctionnement et revient juste à déplacer le problème initial, à savoir comment le cerveau interprète les informations sensorielles. Il est alors nécessaire de faire l'hypothèse de l'existence d'un autre agent intelligent qui lui aussi interprète les informations du premier... et on arrive très rapidement à une régression infinie dont il est impossible de connaître l'origine et le fonctionnement.

Au XX^{ème} siècle, l'avènement de l'informatique basée sur la logique classique et la computation symbolique a mis en évidence que l'on pouvait créer des machines dotées de fonctions spécifiques, comme la possibilité d'automatiser des calculs. L'analogie avec le fonctionnement du cerveau et de la cognition a été très rapide, et de nombreux scientifiques ont alors fait l'hypothèse que, comme un ordinateur, le cerveau s'appuyait sur des symboles, représentations du monde, qu'il manipulait en se basant sur une syntaxe pour raisonner et agir (Changeux, 1983). L'hypothèse sous-jacente est

2. Les considérations physiologiques proposées par Descartes ne sont pas présentées ici et on fait référence à O'Regan (2011) pour plus de détails sur les hypothèses émises par Descartes.

donc que le cerveau est une *machine* qui permet de représenter le monde de façon adéquate, et que ces représentations sont ensuite manipulées. Comme le propose Varela *et al.* (1993)³, le programme cognitiviste peut se définir en plusieurs questions :

Question 1 : Qu'est-ce que la cognition ?

Réponse : Le traitement de l'information : la manipulation de symboles à partir de règles.

Question 2 : Comment cela fonctionne-t-il ?

Réponse : Par n'importe quel dispositif pouvant représenter les éléments physiques discontinus : des symboles. Le système n'interagit qu'avec la forme des symboles (leurs attributs physiques), et non leur sens.

Question 3 : Comment savoir qu'un système cognitif fonctionne de manière appropriée ?

Réponse : Quand les symboles représentent adéquatement quelque aspect du monde réel, et que le traitement de l'information aboutit à une solution efficace du problème soumis au système.

Il est donc clair que le programme cognitiviste dépasse les hypothèses faites par Descartes, le point central de cette approche étant l'existence de représentations. Cette vision de la cognition est encore très soutenue dans la mesure où les outils d'Imagerie par Résonance Magnétique fonctionnelle (IRMf) permettent de mettre en évidence que certaines zones du cerveau semblent avoir une fonction spécifique pour le traitement des informations visuelles ou sonores par exemple. L'approche cognitiviste est donc toujours très utilisée, d'une part car elle fournit un cadre théorique qui semble bien corrélérer les observations neuropsychologiques, et d'autre part car sa terminologie est très intuitive. En effet, parler de représentation mentale par exemple est compréhensible très facilement et réfère bien à l'idée première que l'on se fait de la cognition en tant de bijection du *monde réel* à l'échelle du cerveau. Cependant, cette approche a été et est de plus en plus critiquée, aussi bien au niveau théorique qu'au niveau pratique.

1.1.2 Au commencement était l'action : les approches sensori-motrices de la perception

L'approche cognitiviste postule donc que, pour être viable, la cognition forme une représentation du monde à la façon dont un ordinateur peut mémoriser une photo dans sa mémoire et ensuite effectuer des opérations dessus. Il est cependant impossible de vérifier que les représentations que l'on se fait du monde correspondent bien au monde réel car nous sommes à la fois le sujet et l'objet de nos interrogations. Nous ne pouvons donc pas savoir si la représentation du monde que l'on a correspond bien au monde en soi⁴. De plus, la notion même de représentation pose un problème

3. Francisco J. Varela proposa une alternative au cognitivisme que nous allons détailler dans le paragraphe qui suit. Dans son livre *L'inscription corporelle de l'esprit*, Varela décrit les différentes approches de la perception, et en particulier l'approche cognitiviste, qui l'ont amenées à en proposer une alternative. Nous ne présenterons pas l'approche connexionniste évoquée par Varela. Cette approche propose une alternative à la préexistence des symboles cognitivistes qui met en avant le fonctionnement en réseau auto-organisé du cerveau et la notion d'émergence. Cette approche est intéressante mais ne propose pas un changement de paradigme aussi radical que les théories motrices de la perception que nous allons présenter ensuite. Pour une description détaillée de cette approche, nous renvoyons le lecteur au livre *L'inscription corporelle de l'esprit* de Varela *et al.* (1993) ou à l'ouvrage *Invitation aux sciences cognitives* de Varela (1996) pour une présentation plus vulgarisée.

4. On peut prendre l'exemple de la perception de la couleur, les pathologies comme le daltonisme montrent bien que la perception de la couleur est subjective et qu'elle est dépendante de la physiologie du système optique.

terminologique et ontologique. En effet, Varela *et al.* (1993) proposent de préciser la notion de représentation et affirment que cette notion ne peut exister que dans son sens *faible*. Ainsi, la représentation mentale d'un stimulus signifie que *la cognition correspond toujours à l'interprétation – ou à la représentation – de l'état du monde* (Varela *et al.*, 1993). Cette affirmation est presque tautologique car c'est évidemment un des rôles de la cognition : interpréter les informations émanant de nos différents sens pour les utiliser et interagir avec. Cependant, dès l'instant où l'on dépasse ce sens faible de la représentation et que l'on fait l'hypothèse d'une part que le monde est représenté dans le cerveau à une échelle différente, et d'autre part que le cerveau se sert de ces représentations pour raisonner et agir, les implications ontologiques sur le fonctionnement des processus cognitifs sont beaucoup plus fortes. En effet, cela suppose que notre système cognitif possède, dès la naissance, une sorte de carte du monde pré-existante dans lequel on existe, et que l'on évolue et apprend en développant cette carte et les outils nécessaires pour la manipuler. Dans *Le sens du mouvement*, Berthoz (1997) préfère même s'affranchir dans la mesure du possible du terme représentation qui selon lui nous leurre sur la nature même de la cognition, le terme représentation étant selon lui *trop contaminé par l'idée d'image visuelle* et amènera nécessairement à un dualisme entre l'esprit et l'environnement, dualisme largement remis en question comme nous allons le voir.

Afin de dépasser cette vision constructiviste où le sujet construit une représentation du monde à partir de représentations internes et des stimulations qui arrivent à ses organes sensoriels, différentes approches ont proposé une alternative qui intègre le corps et la motricité dans la formation des percepts cognitifs. En effet, la séparation entre le monde réel et la cognition peut apparaître très arbitraire. La première approche⁵ intéressante à présenter est l'approche proposée par J. J. Gibson, également appelée approche écologique de la perception. Gibson (1966) propose que l'on étudie la cognition comme le moyen des êtres vivants d'interagir avec l'environnement dans lequel ils évoluent. Cela signifie que la cognition extrait les caractéristiques de l'environnement par le biais des capacités sensori-motrices des individus. Les informations extraites d'un flux sensoriel sont donc les informations utiles pour avoir un comportement adapté dans l'environnement. Gibson fait ainsi l'hypothèse que l'environnement et les flux sensoriels, captés par un être vivant contiennent des informations invariantes qui lui permettent d'interagir de façon adaptée. Cette hypothèse a été formalisée par Gibson (1979) avec la notion d'*affordance* qui correspond aux possibilités offertes par l'environnement aux êtres vivants par rapport à leurs capacités sensori-motrices. Shepard (1984) propose une métaphore élégante de la façon dont pourrait fonctionner cette boucle sensori-motrice, il suggère qu'à l'image d'une corde de piano notre cerveau *résonne* sur des modes excités par les flux sensoriels, les modes sur lesquels il résonne étant définis par la structure du cerveau et au grès de l'expérience des sujets. Varela *et al.* (1993) fournissent un exemple intéressant permettant de clarifier la notion d'*affordance*, ils prennent l'exemple des arbres, possibilité offerte par l'environnement permettant à certains animaux de grimper dessus car ils en ont les capacités motrices, l'arbre *afforde* donc l'action d'escalader. De plus, Gibson nie farouchement le concept de représentation et soutient que la perception et l'action ne nécessitent aucun traitement symbolique. L'approche écologique est ainsi parfois appelée *perception directe* (Michaels et Carello, 1981) car elle postule que le monde est perçu directement (sans interprétation) et que l'action est directement guidée par les informations invariantes des flux sensoriels.

Gibson va donc à l'encontre des théories cognitivistes en replaçant l'action au coeur du proces-

5. De nombreuses théories intégrant la motricité au sein des processus cognitifs ont été élaborées et nous n'en présentons qu'une petite partie, pour revue plus détaillée, nous renvoyons au chapitre premier de l'ouvrage *Le sens du mouvement* d'Alain Berthoz (1997) qui les présente de façon plus exhaustive.

sus de perception, et affirme que la perception n'est pas possible sans interaction sensori-motrice avec l'environnement. Cependant, Gibson postule que l'environnement est complètement indépendant du sujet percevant. Varela *et al.* (1993) proposent quant à eux de dépasser cette dualité entre le sujet et l'environnement en affirmant que c'est le sujet, par ses interactions sensori-motrices, qui fait émerger l'environnement. La différence avec l'approche écologique est subtile mais mérite d'être précisée. Quand Gibson place le sujet au sein de l'environnement, Varela place l'environnement et le sujet au même niveau. Dit autrement, le sujet et l'environnement perçus sont le résultat d'un couplage structurel entre le cerveau, le corps et l'environnement. Selon l'approche *énactive* proposée par Varela, l'environnement que l'on perçoit n'existe pas indépendamment du sujet, et les deux entités co-évoluent à l'échelle du sujet, mais également à l'échelle de l'espèce. Il est donc impossible de comprendre la cognition en ignorant l'histoire du couplage entre l'espèce (dans nos études l'espèce humaine) et son environnement (dans nos études le monde), mais également en ignorant le couplage entre un sujet (un individu de l'espèce humaine dans nos études) et son histoire individuelle (son expérience depuis sa naissance). Pour définir le programme de l'approche dite de l'énaction⁶, Varela *et al.* (1993) ainsi proposent de répondre aux mêmes questions que celles posées pour définir l'approche cognitive :

Question 1 : Qu'est que la cognition ?

Réponse : L'action productive : l'historique du couplage structurel qui énonce (faire-émerger) un monde.

Question 2 : Comment cela fonctionne-t-il ?

Réponse : Par l'entremise d'un réseau d'éléments inter-connectés, capable de subir des changements structuraux au cours d'un historique non interrompu.

Question 3 : Comment savoir qu'un système cognitif fonctionne de manière appropriée ?

Réponse : Quand il s'adjoint à un monde de signification préexistant, en continu développement (comme c'est le cas des petits de toutes les espèces) ou qu'il en forme un nouveau (comme cela arrive dans l'histoire de l'évolution).

Les enjeux du débat sur la viabilité des programmes de l'approche écologique et éactive par rapport à celui de l'approche cognitive dépassent largement le cadre et l'ambition de cette thèse. Cependant, il nous a semblé essentiel de présenter ces différents aspects car ils fournissent un cadre plus formel pour aborder la perception sonore des sons du quotidien et plus spécifiquement le problème du geste imaginé quand on entend un son produit par ce geste. En effet, l'approche cognitive ne nous fournit aucune hypothèse sur les gestes que l'on peut imaginer à partir d'un son. A l'inverse, l'approche éactive et l'approche écologique nous fournissent elles des hypothèses fortes sur les possibilités offertes par l'environnement et les compétences sensori-motrices des êtres humains mises en jeu dans la formation des représentations (au sens faible !) que l'on peut se faire d'un geste à partir d'un son. Dans une perspective épistémologique plus large, ces deux approches obligent le chercheur à intégrer sa subjectivité au sein même de son programme recherche. On peut mentionner des travaux de Devereux et Sinaceur (1967) issus de recherches en psychanalyse qui proposent que l'observateur ne peut comprendre le sujet observé, ici la cognition, que si il intègre sa propre subjectivité

6. Le verbe anglais *to enact* peut être traduit par *faire émerger* qui traduit bien l'idée que le sujet est moteur de sa perception par son interaction avec le monde sensible.

au sein de son analyse, ici les contraintes sensori-motrices. Pour comprendre la cognition, les perspectives écologiques et éenactives nous incitent ainsi à nous placer à un niveau où le monde est perçu par le biais de notre corps et de nos sens, sources de subjectivité en raison de leurs spécificités structurelles. Ainsi, si l'on comprend la structure même de la perception, nous devrions pouvoir (mieux) comprendre la façon dont fonctionne la cognition.

Avec l'évolution des méthodes d'imagerie cérébrale et les connaissances récentes issues des neurosciences cognitives, il est désormais possible de confronter des données issues de ces approches psychologiques avec des mesures objectives permettant de trancher certains aspects incompatibles entre les approches cognitivistes et les approches sensori-motrices intégrant l'action au coeur des processus perceptifs. Une autre approche, appelée approche idéomotrice (Greenwald, 1970; Prinz, 1997; Hommel *et al.*, 2001; Norman, 2002) (aussi discutée par Young *et al.* (2013) dans le cas de l'audition) propose une réunification de ces approches *a priori* incompatibles au regard de nouveaux modèles neuropsychologiques (Cisek et Kalaska, 2010). Ces aspects ne seront pas présentés dans l'état de l'art car ils ne sont pas nécessaires pour aborder les enjeux de cette thèse, cependant ils seront abordés dans l'introduction et la discussion du chapitre 2. Le fait que les processus moteurs interviennent directement dans les processus de perception a donc été mis en évidence dans diverses études issues des neurosciences cognitives que nous allons rappeler dans le paragraphe qui suit.

1.1.3 La perception : une action simulée

Pour introduire ces études, nous allons juste évoquer succinctement une approche qui n'a pas été mentionnée précédemment, l'approche proposée par Kevin O'Regan (2011)⁷. Kevin O'Regan s'est principalement intéressé à la vision et l'approche qu'il propose est assez similaire *en première approximation* avec les approches présentées dans la section précédente. Elle propose en particulier que *voir, ce n'est pas se faire une représentation du monde extérieur par le biais du système visuel*. Dans O'Regan (1992); O'Regan et Noë (2001); O'Regan (2011), les auteurs proposent que, *voir, c'est envisager les choses que l'on peut faire avec ce qui est perçu*, et proposent également de considérer le monde comme *une mémoire externe* dans lequel le système perceptif va extraire des informations par le biais des *contingences sensori-motrices* du sujet percevant. Par exemple, voir un verre de vin ne correspond pas au fait d'avoir une image du verre de vin dans la tête, mais c'est imaginer l'ensemble des actions que l'on peut faire avec, par exemple le remplir ou le boire. Berthoz (1997) et Jeannerod (2011) proposent une idée similaire en suggérant que percevoir c'est simuler l'action.

Cette relation intrinsèque entre perception et action a été validée par des études issues de recherches en neurosciences cognitives chez l'animal. Rizzolatti *et al.* (1996); Gallese *et al.* (1996) ont par exemple mis en évidence que des neurones du cortex ventral prémoteur de singes, appelés désormais neurones miroirs, déchargeaient aussi bien quand le singe réalisait une action que lorsqu'il la voyait seulement sans pour autant la réaliser. Ces observations ont été également faites dans le cas de la modalité auditive. Kohler *et al.* (2002) ont montré que des neurones du cortex prémoteur déchargeaient aussi bien quand les singes entendaient les sons d'une action spécifique, par exemple casser une cacahuète, que quand ils la réalisaient. Ces résultats séminaux confortent donc les hypothèses formulées dans les approches sensori-motrices, stipulant que l'action et la perception sont bien des phénomènes complètement intriqués à très bas niveau dans l'organisation des processus cérébraux.

Des corrélats entre perception et action ont également été mis en évidence chez l'homme, et en particulier dans des tâches d'imageries motrices. L'imagerie motrice, et plus généralement l'imagerie

7. On peut également mentionner l'ouvrage *Action in perception* d'Alva Noë (2004).

mentale, est un domaine étudié depuis très longtemps en psychologie expérimentale. Dans ce type d'expérience, il est demandé aux sujets de s'imaginer une situation ou une action. Les expérimentateurs évaluent alors par différents moyens les caractéristiques de ce que le sujet a imaginé. Une méthode très utilisée est la chronométrie mentale. Cette méthode consiste à mesurer le temps mis par le sujet pour réaliser la tâche d'imagination. Il est ensuite possible de déterminer des corrélations éventuelles entre le temps mis par le sujet pour s'imaginer la tâche et la durée qu'il faudrait pour la réaliser réellement. Decety *et al.* (1989) ont par exemple montré que le temps mis pour s'imaginer marcher jusqu'à une cible visuelle était comparable au temps réellement mis pour atteindre cette cible en marchant. Cela suggère donc que les processus mentaux impliqués dans ces différentes tâches sont régis par les mêmes processus cérébraux. Dans le cas des mouvements, Parsons (1994) a montré que le temps mis pour s'imaginer une posture de la main et le temps mis pour mettre sa main dans la même position sont comparables quand la posture de la main est facile à réaliser, et qu'il est plus court, mais proportionnel, dans le cas où la posture de la main est plus difficile à réaliser. Ce dernier point suggérant donc que l'on s' imagine la cinématique du mouvement. Dans le cas spécifique des mouvements graphiques qui nous intéressent plus particulièrement dans cette thèse, Decety et Michel (1989) ont montré un résultat similaire. Des sujets devaient s'imaginer dessiner un cube ou écrire une phrase donnée. Ils ont alors observé que le temps mis pour s'imaginer ces deux tâches motrices était comparable au temps que les mêmes sujets mettaient pour dessiner réellement le cube ou écrire la phrase. Une étude récente de Papaxanthis *et al.* (2012) a précisé ce résultat. Des sujets devaient s'imaginer en train de dessiner différentes formes de même longueur mais dont la répartition de courbure variait le long de la trajectoire en fonction des formes. Ils ont pu mettre en évidence que le temps mis pour s'imaginer dessiner les formes était proportionnel à la répartition de courbure le long de la forme : plus la forme présentait des zones avec des grandes courbures, plus le temps mis pour s'imaginer la dessiner était grand. Les résultats de Papaxanthis *et al.* (2012) sont donc cohérents avec la loi en puissance $1/3$. En effet, plus une forme présente des portions avec des fortes courbures, plus la vitesse moyenne est faible, et donc plus la durée pour la dessiner est grande. Cela montre donc que l'imagination d'une tâche motrice a des caractéristiques dynamiques correspondant aux caractéristiques dynamiques réelles. Au niveau neuropsychologique, des études d'imagerie fonctionnelle ont également montré que les zones corticales activées quand une personne voyait une forme manuscrite étaient les mêmes que quand elle la produisait (Longcamp *et al.*, 2003, 2006, 2008; James et Gauthier, 2006).

Au niveau de l'audition, qui est le sujet d'étude principal de cette thèse, la relation entre perception et motricité a été mise en évidence dans le cas très particulier de la perception de la parole. La parole est en effet un acte moteur, et les mouvements réalisés sont souvent appelés gestes phona-toires. De nombreuses expériences ont mis en évidence que nous ne percevons pas les signaux de paroles comme de *simples* signaux acoustiques mais comme des actes moteurs potentiels (Liberman *et al.*, 1967; Liberman et Mattingly, 1985). De plus il a été montré que nous avons appris à fusionner les sons produits par les objets et les actions qui les ont générés (Hommel, 2004). Des tâches d'imagerie mentales *sonores* ont de plus confirmé que lorsque l'on s'imaginait un son, un morceau de musique par exemple, le cortex auditif s'activait de la même façon que lorsque l'on entendait réellement un son ((Zatorre et Halpern, 2005) pour une revue rapide). Et dans le cas des musiciens, il a même été montré qu'entendre le son de l'instrument qu'ils savent jouer activait les aires motrices activées quand ils jouent réellement de leur instrument, chez les pianistes professionnels par exemple (Haueisen et Knösche, 2001; Bangert *et al.*, 2006).

Toutes ces études suggèrent donc que les processus cognitifs et les processus neuropsycholo-

giques impliqués dans la perception, la simulation, ou la production d'une action sont communs et en particulier du point de vue de l'audition.

Synthèse

On a donc vu dans cette section que la perception et l'action étaient complètement reliées : percevoir, c'est s'imaginer réaliser l'action sous-jacente. On peut résumer en quelques points les idées les plus importantes :

1. pour étudier la cognition, et en particulier la cognition auditive, il est nécessaire de tenir compte des caractéristiques de nos contraintes sensori-motrices ;
2. les caractéristiques des processus cognitifs impliqués dans la perception ou l'imagination d'une action sont en partie similaires à ceux mis en jeu dans la production de la même action ;
3. au regard des ambiguïtés ontologiques impliquées par les termes *représentation mentale* ou *image mentale*, il est nécessaire de préciser que dans ce manuscrit nous ne les utiliserons que dans leurs acceptation faible comme présenté précédemment. Ce sens n'implique donc aucune hypothèse sur la nature des processus cognitifs sous-jacents. Cette terminologie sera donc éventuellement employée dans la suite de la thèse uniquement pour caractériser l'action d'imaginer à partir d'un stimulus sonore et les caractéristiques de ce qui est imaginé. Nous ne postulons donc en aucun cas qu'une image mentale est reproduite et représentée quelque part dans le cerveau.

1.2 Approche écologique de la perception auditive

Les différentes approches présentées dans la section précédente, surtout autour d'exemples issus de la modalité visuelle, nous fournissent des pistes intéressantes pour envisager comment aborder le problème du geste que l'on imagine en entendant un son de frottement produit par une personne qui dessine. La section suivante va s'attacher à présenter le cadre précis dans lequel nous allons aborder le problème de la perception de ces sons. Cette approche se base sur l'approche écologique de Gibson en adaptant le concept d'invariant d'un flux sensoriel à la modalité auditive (McAdams, 1993).

En s'appuyant sur l'approche proposée par Gibson (1966), Warren et Verbrugge (1984); McAdams (1993) proposent de définir deux types d'invariants reliés à la perception et la reconnaissance d'un événement sonore, les *invariants structuraux* caractérisant la nature de la source sonore (par exemple le matériau ou la forme d'un objet impacté par exemple), et les *invariants transformationnels* caractérisant l'action qui est effectuée sur l'objet (par exemple frotter, frapper ou rebondir pour un objet solide). Les invariants acoustiques sont donc les propriétés acoustiques contenues dans le signal sonore qui transmettent l'information utilisée par le cerveau pour reconnaître un événement sonore. Cette discrimination entre l'action et l'objet est essentielle et sera utilisée dans notre paradigme d'étude pour synthétiser des stimuli en considérant séparément ces deux aspects de la perception d'un événement sonore.

Dans cette section, nous allons présenter une revue des études s'étant intéressées à ces deux types d'invariants dans le cas des sons produits par des interactions solides, correspondant à la catégorie intégrant les sons de frottement que nous allons étudier plus spécifiquement dans la thèse. D'une façon très générale, de nombreuses études ont mis en évidence que nous étions capables d'identifier certaines propriétés d'un événement sonore sans pour autant avoir d'informations visuelles (Lutfi, 2007). Différentes études se sont intéressées plus particulièrement à comprendre comment nous catégorisons les sons d'environnement (i.e. *tous les sons autres que les sons musicaux et les signaux de parole*

(Gygi et Shafiro, 2006)). On peut en particulier mentionner Gaver (1993a) qui a suggéré que nous catégorisons les sons de l'environnement par rapport une taxonomie spécifique caractérisée par le type d'interaction qui a produit le son : solide, liquide et aérodynamique. Houix *et al.* (2012) ont récemment validé cette taxonomie du point de vue perceptif dans des tâches de classification libre lors desquelles les sujets devaient classer 60 sons en plusieurs catégories. Les sujets ont alors classé naturellement les sons en fonction des différentes catégories proposées par Gaver (1993a).

1.2.1 Perception auditive des objets : les *invariants structuraux*

Beaucoup d'études se sont intéressées aux sons d'impacts (on peut mentionner en particulier (Giordano *et al.*, 2010; Grassi, 2005; Grassi *et al.*, 2013; Kunkler-Peck, 2000; McAdams *et al.*, 2004; Aramaki et Kronland-Martinet, 2006; Aramaki *et al.*, 2011)) qui permettent de se focaliser sur les propriétés acoustiques et perceptives de l'objet impacté. A partir des informations acoustiques contenues dans le son généré, il est ainsi possible de reconnaître différentes caractéristiques de l'objet, telles que son matériau ou sa forme. Ces études ont par exemple mis en évidence que les caractéristiques physiques liées à la dispersion et la dissipation sont essentielles pour la reconnaissance d'une source sonore. De plus, il a été mis en évidence que les sons d'impact contiennent suffisamment d'information pour permettre la discrimination du matériau des objets impactés (Aramaki *et al.*, 2011; Klatzky *et al.*, 2000; Wildes et Richards, 1988), et dans une certaine mesure reconnaître leurs formes (Carello *et al.*, 1998; Lakatos *et al.*, 1997). En particulier, Giordano et McAdams (2006) ont évalué l'influence de la taille d'un objet sur la reconnaissance du matériau d'un objet impacté et ont ainsi mis en évidence des descripteurs acoustiques robustes permettant d'expliquer l'identification du matériau. Carello *et al.* (1998) ont eux montré que l'on pouvait reconnaître la longueur d'un objet à partir d'un son d'impact, tandis que Lakatos *et al.* (1997) ont eux montré que l'on pouvait en déduire la forme si elle était assez simple.

Pour la plupart des objets du quotidien, un son d'impact peut être modélisé comme une somme de N sinusoides exponentiellement amorties (Aramaki *et al.*, 2011) :

$$s(t) = \sum_{n=0}^{N-1} A_n \sin(2\pi f_n t + \phi_n) e^{-\alpha_n t} \quad (1.1)$$

où A_n correspond à l'amplitude de chaque mode, f_n la fréquence du mode n de l'objet, ϕ_n la phase du mode n , et α_n l'amortissement du mode n d'une loi d'amortissement proposée par les auteurs et définie par la formule $\alpha(f_n) = e^{\alpha_G + \alpha_R 2\pi f_n}$ avec α_G , l'amortissement global et α_R , l'amortissement relatif (permettant de rendre compte de la dépendance fréquentielle de l'amortissement). Au-delà de montrer que nous sommes capables de discriminer différents matériaux, certaines études se sont intéressées à caractériser l'évolution de cette perception du matériau en fonction des différents paramètres de l'équation 1.1. La perception du matériau est principalement régie par les valeurs d'amortissement qui sont fonction de la fréquence. En réalisant des tests de catégorisation, Aramaki *et al.* (2011) ont cartographié un espace sonore du matériau perçu d'un objet impacté (bois, métal, verre) en fonction des deux paramètres. Un corrélat acoustique de la perception de la forme concerne la fréquence fondamentale f_0 du premier mode de l'objet, plus f_0 sera élevée plus l'objet sera perçu petit et réciproquement (Lakatos *et al.*, 1997; Kunkler-Peck, 2000; Aramaki *et al.*, 2010).

Dans le cadre de la perspective écologique et en adoptant la taxonomie des invariants, on peut dire finalement que l'amortissement, plus spécifiquement α_R et α_G , et la fréquence fondamentale f_0 , sont des exemples d'invariants structuraux en partie responsables de la reconnaissance du matériau et de la taille des objets solides. Nous allons maintenant nous intéresser aux invariants transforma-

tionnels qui concernent la perception auditive de l'action qui est réalisée sur un objet.

1.2.2 Perception auditive des actions : les *invariants transformationnels*

Avant de présenter les études spécifiques sur les invariants transformationnels, il est intéressant de noter qu'une étude a montré que la reconnaissance auditive des actions effectuées sur un objet est plus rapide que la reconnaissance des propriétés de l'objet. On pourrait dire que l'on extrait plus vite les invariants transformationnels que les invariants structuraux. Dans des expériences de catégorisation, Lemaitre et Heller (2012) ont en effet montré que des sujets catégorisaient significativement plus rapidement les actions qui ont produit un son que les propriétés physiques des objets (e.g. le matériau) sur lesquels ont été réalisées les actions. Ce résultat suggère donc que notre système auditif est plus adapté pour reconnaître l'action qui a mis en vibration l'objet que les propriétés de l'objet lui-même.

Concernant les invariants transformationnels, Warren et Verbrugge (1984) ont montré qu'on pouvait prédire si un objet allait rebondir ou se casser juste en se basant sur le rythme d'une série d'impacts contenus dans le son produit. Li *et al.* (1991) ont eux montré que le genre d'une personne pouvait être reconnu juste en écoutant les sons produits par ses bruits de pas. La reconnaissance du genre féminin est par exemple caractérisée par la perception des différences d'amplitudes entre les pics du spectre et la présence plus ou moins importante de hautes fréquences. Ces deux études sont très intéressantes car elles montrent que du point de vue perceptif, le pattern temporel d'une suite d'impacts est vraiment caractéristique de l'action qui a été réalisée.

Plus récemment, Conan *et al.* (2014b) ont mis en évidence que les interactions continues comme le frottement, le grattement, ou le roulement, pouvaient être reconnues sur la base des caractéristiques temporelles des suites d'impacts d'un excitateur sur un résonateur. En particulier, Conan *et al.* (2014a) ont montré que les sons de roulement se distinguaient perceptivement des sons de frottement et des sons de grattement par la corrélation entre la séparation temporelle entre deux impacts et leurs amplitudes respectives. Le cas des sons de frottement qui nous intéresse particulièrement dans notre étude sera abordé plus en détail dans une section ultérieure.

Enfin, il est intéressant de constater que l'invariant structurel d'un objet peut être reconnu à travers les différents invariants transformationnels qui lui sont associés. Ainsi, une balle peut être reconnue à travers différentes actions comme le rebond ou le roulement. De la même façon, quelques études ont montré que le son d'une balle qui roule permettait de reconnaître sa taille et sa vitesse (Houben *et al.*, 2004, 2005).

1.2.3 Perception auditive des mouvements

Dans la suite, nous allons nous intéresser plus particulièrement à la perception des sons évoquant un mouvement. En adoptant la taxonomie des invariants, on peut dire qu'un des objectifs des travaux qui vont être présentés est de mettre en évidence les invariants transformationnels acoustiques reliés aux gestes humains. De nombreuses études se sont déjà intéressées à la perception du mouvement par le biais de la modalité auditive. Afin de bien spécifier notre approche, il convient de préciser rapidement les différences entre ces études et la façon dont nous allons aborder la relation entre des mouvements spécifiques, les gestes humains, et la perception sonore.

La perception du mouvement d'une source sonore a surtout été étudiée du point de vue psychoacoustique. Harris et Sergeant (1971) ont par exemple déterminé le plus petit mouvement d'angle audible d'une source sonore en mouvement pour des bruits large bande, Carlile et Best (2002) et Grantham (1986) ont eux déterminé les seuils de discrimination de la vitesse angulaire d'une source

sonore en mouvement. Dans ces études, les stimuli utilisés étaient des sons spatialisés par différentes méthodes (utilisation des HRTF par exemple).

Cependant, le mouvement d'une source sonore peut également avoir des conséquences sur les variations intrinsèques des descripteurs acoustiques du son, l'exemple le plus connu est le cas de l'effet doppler. Quand une voiture se déplace, la source sonore se déplace, mais une variation intrinsèque du son est également perçue par un auditeur fixe. Dans le cas d'une source monochromatique (i.e. ne contenant qu'une seule fréquence), la fréquence perçue par l'auditeur varie en fonction de la vitesse de la source. En première approximation, la fréquence perçue correspond à la fréquence de la source multipliée par un facteur $(1 - \frac{v_s}{c})$, avec v_s la vitesse de la source et c la vitesse du son dans l'air. L'intensité du son varie également en fonction de la vitesse de la source. En première approximation, la pression de l'onde acoustique arrivant aux oreilles de l'auditeur est multipliée par un facteur $(1 - \frac{v_s}{c})^2$. On peut ainsi reproduire ce type d'effet très facilement sans pour autant utiliser une source stéréo ou spatialisée. Un son monophonique peut être synthétisé en utilisant ce type de relation entre la vitesse et la fréquence (Chowning, 1971; Kronland-Martinet et Voinier, 2008) et il évoquera le mouvement de la source de façon très réaliste. Cet exemple montre donc qu'un mouvement peut être perçu par la modalité auditive de deux façons, soit par les variations d'indices spatiaux – principalement la différence de temps d'arrivée (ITD : Interaural Time Difference), et d'intensité (ILD : Interaural Level Difference) du son due à la distance entre les deux oreilles – soit par des variations des propriétés intrinsèques au son – sa composition fréquentielle, son intensité, et son timbre.

Dans les études qui vont être présentées par la suite, les mouvements évoqués par les sons ne concernent que la seconde situation, i.e. des sons monophoniques qui varient par leurs propriétés intrinsèques. On préfère insister sur ce point afin qu'il n'y ait pas d'ambiguïté dans la suite du manuscrit, nous nous intéresserons donc à la perception du mouvement évoqué par les variations intrinsèques du son et non à la perception du mouvement réel d'une source sonore. Nous parlerons le plus souvent de la dynamique du son pour caractériser le mouvement évoqué par le son. En effet, ce qui est perçu, ce sont les conséquences du mouvement physique sur les caractéristiques acoustiques du signal. Les variations intrinsèques d'un son sont caractérisées par différents descripteurs acoustiques, à savoir le barycentre spectral et l'intensité ainsi que ceux relatifs au timbre (Von Helmholtz, 1912). Le timbre définit classiquement toutes les propriétés acoustiques autre que la fréquence, l'intensité, la durée ou encore la réverbération, d'un son permettant de discriminer deux sources sonores.

Des études ont tenté de définir plus précisément les différentes dimensions du timbre (Berger, 1964), et en particulier dans le cas des sons produits par les instruments de musique (Grey et Gordon, 1978; McAdams *et al.*, 1995). Dans McAdams *et al.* (1995), trois descripteurs de timbre semblent bien décrire les dimensions de l'espace perceptif associé : l'attaque du son, le flux spectral et la brillance. Les deux dernières nous intéressent plus particulièrement car les variations de timbre des sons de frottement y sont énormément reliées. Le flux spectral caractérise la façon dont évolue la répartition spectrale d'énergie au cours du temps, et la brillance est une dimension du timbre reliée à l'étalement spectral d'un son autour de son centre de gravité spectral. Si on considère un son qui évolue au cours du temps, des variations de ces descripteurs peuvent être fortement impliquées dans l'évocation des mouvements. Dans des études très générales sur la perception et la synthèse de sons évoquant des mouvements, Merer *et al.* (2013, 2008); Merer (2011) ont mis en évidence que des variations tempo-

relles de ces descripteurs pour des sons abstraits⁸ caractérisent bien des sons évoquant des mouvements. Ils ont d'ailleurs utilisé ces constatations pour créer un synthétiseur de sons contrôlé par une description sémantique du mouvement évoqué par le son. Ils ont ainsi mis en évidence que l'on pouvait distinguer plusieurs grandes classes de mouvements évoqués par les sons, par exemple "son qui tourne" ou encore "son qui passe". Cette approche de la perception du mouvement évoqué par les sons a également été abordée par le Laboratoire de Musique et Informatique de Marseille (MIM) qui s'est intéressé particulièrement à la description des sons en terme d'Unité Sémiotique Temporelle (UST). On peut définir les UST comme l'ensemble des *figures* de base décrivant l'évolution temporelle d'un son. Frémiot *et al.* (1996) ont ainsi sélectionné 19 figures de référence, chacune d'entre elle étant associée à une description sémantique. Un certain nombre de ces figures réfèrent directement à des notions de mouvement (par exemple "chute", "trajectoire inexorable", "qui avance"). Les UST mettent donc en évidence le rôle prépondérant du mouvement dans la perception des sons et en particulier des sons musicaux, ces éléments de base pouvant ensuite être réutilisés au cours du processus de composition musicale pour évoquer des mouvements particuliers (dans la musique électroacoustique par exemple). En ce qui concerne les sons générés par des interactions continues comme les sons de frottement, on a déjà mentionné les études d'Houben *et al.* (2004, 2005) qui ont mis en évidence que l'on pouvait reconnaître, dans une certaine mesure, la vitesse d'une bille qui roule juste à partir des variations de timbre du son produit.

Synthèse

Cette section a donc présenté une revue des connaissances sur la perception des sons du quotidien, on peut résumer ces connaissances en quelques points :

1. le système auditif permet de reconnaître des événements sans pour autant utiliser d'informations visuelles ;
2. en adoptant le paradigme proposé par l'approche écologique, on peut considérer que la perception d'un événement sonore comme un son de frottement est caractérisée par un invariant structurel, lié à la perception de l'objet, et un invariant transformationnel, lié à la perception de l'action réalisée sur l'objet ;
3. un son peut évoquer un mouvement sans pour autant être spatialisé. Les variations de la structure interne d'un son, caractérisées par différents descripteurs acoustiques, permettent en effet d'évoquer des mouvements de façon très réaliste dans le cas des sons produits par les interactions continues comme les frottements.

1.3 Production motrice des mouvements biologiques

Les deux sections précédentes ont donc présenté des approches psychologiques très générales concernant la perception et la cognition, et plus particulièrement l'approche écologique de la perception sonore. Une information essentielle à retenir pour les travaux qui vont être présentés dans cette thèse est que pour étudier la perception des sons produits par les gestes graphiques, il est nécessaire de se concentrer sur les caractéristiques de ces mouvements. En effet, notre système auditif semble se focaliser sur les informations les plus caractéristiques des actions, i.e. les invariants transformationnels, pour les reconnaître. Ainsi, pour aborder le problème des relations entre sons et gestes, nous avons également besoin de connaissances issues des sciences du mouvement pour comprendre la façon dont nous produisons nos mouvements.

8. Sons n'évoquant pas de source sonore facilement identifiable.

1.3.1 Invariants cinématiques des mouvements biologiques

Quand une personne dessine une forme ou écrit de façon spontanée, les caractéristiques de son mouvement sont reliées aux caractéristiques géométriques de la forme qu'elle dessine, en particulier sa taille et ses variations de courbure.

Isochronie.

L'isochronie est une propriété qui caractérise le temps qu'une personne met pour dessiner une forme. Elle stipule que le temps mis pour dessiner une forme est globalement constant quelle que soit sa taille. Dit autrement, l'isochronie révèle que le temps moyen mis pour dessiner une forme est constant quelle que soit la taille de la forme dessinée. Cette propriété a été mise en évidence dans des travaux très anciens sur l'écriture manuscrite de Binet et Courtier (1893), et a ensuite été confirmée dans de nombreuses études. On peut mentionner en particulier une étude de Viviani et McCollum (1983) qui l'ont montrée dans le cas de formes périodiques comme des cercles ou des ellipses, formes qui seront largement utilisées dans les expériences menées au cours de cette thèse. De plus, dans une expérience d'imagerie motrice, Decety et Michel (1989) ont mis en évidence que le temps mis pour s'imaginer une forme était le même que celui mis pour la dessiner, et que le principe d'isochronie était ainsi également respecté dans cette tâche n'impliquant aucun mouvement. Cela suggère que l'isochronie est intégrée également du point de vue perceptif.

La loi en puissance 1/3.

L'isochronie met donc en évidence une invariance globale sur la durée de production d'une forme graphique, cependant, elle ne dit rien sur la façon dont varie la vitesse de production au cours du tracé d'une forme. En plus de s'intéresser à l'isochronie, Binet et Courtier (1893) ont également constaté que l'on ralentissait dans les parties les plus courbées et que l'on accélérât dans les parties les plus plates d'une trajectoire (on peut faire référence aux travaux de Jack (1895)). Ces observations holistiques ont été précisées avec le développement de technologies comme les tablettes graphiques permettant d'enregistrer précisément les mouvements au cours du temps. Les travaux séminaux qui font référence dans le cadre de l'étude des mouvements graphiques sont ceux de Paolo Viviani et de ses collègues qui ont précisé les observations précédentes en s'intéressant aux covariations de la vitesse et de la courbure au cours d'un mouvement graphique. Ils ont en particulier montré que les variations temporelles de la vitesse angulaire $A(t)$ d'un mouvement étaient reliées aux variations temporelles de la courbure $C(t)$ de la forme dessinée au cours du temps par une loi de puissance $2/3$ (Lacquaniti *et al.*, 1983; Viviani et Terzuolo, 1982; Viviani, 1986) :

$$A(t) = KC(t)^{\frac{2}{3}} \quad (1.2)$$

où K est une constante appelée gain de vitesse reliée à la vitesse moyenne du mouvement. Cette loi peut également s'exprimer en fonction de la vitesse tangentielle $v_t(t)$ du mouvement :

$$v_t(t) = KC(t)^{-\frac{1}{3}} \quad (1.3)$$

Cette relation formalise donc les observations de Binet et Courtier (1893) et Jack (1895), et montre donc que l'on accélère dans les parties les plus plates et que l'on ralentit dans les parties les plus courbées d'une trajectoire.

Cette relation a été largement étudiée par la suite et a été validée dans de nombreuses situations. Viviani et Flash (1995) ont en particulier montré que, comme l'isochronie, cette invariance cinématique est respectée quelle que soit l'amplitude du mouvement. Cependant, Teulings *et al.* (1986) et

Plamondon (1998) ont également montré que cette loi pouvait être violée dès l'instant où la forme dessinée contenait des points de discontinuité (i.e. des points de rebroussement).

Cette invariance est très intéressante, et l'on pourra ainsi se demander s'il est possible de reconnaître cette loi à travers les variations de timbre des sons de frottement.

1.3.2 Dynamique et coordination des mouvements biologiques

Il semble donc que la loi en puissance $1/3$ soit une signature robuste des mouvements produits par les humains lorsqu'ils dessinent. Cependant, cette relation n'offre qu'une description des phénomènes et ne fournit pas d'hypothèse sur son origine.

Des études se sont également intéressées aux relations entre les caractéristiques mécaniques des mouvements et leurs conséquences au niveau dynamique. Les objectifs de cette thèse n'étant pas focalisés sur la production des mouvements mais sur leur perception auditive, nous nous attacherons à rappeler ici les concepts et connaissances nécessaires pour la compréhension des études que nous avons menées sur la perception auditive des sons produits par les mouvements graphiques.

Afin de présenter l'approche dynamique de la production des mouvements il est nécessaire d'en préciser rapidement ses origines. L'isochronie et la loi en puissance $1/3$ décrivent la production des mouvements en terme d'invariance. Les modèles sous-jacents font donc l'hypothèse que le contrôle du mouvement est réalisé d'une part par un contrôle proactif, c'est-à-dire par l'activation de schémas moteurs en mémoire contrôlant les propriétés d'invariances comme l'isochronie et la loi en puissance $1/3$, mais également pas le biais d'un contrôle rétroactif réalisé à partir d'informations visuelles et haptiques. Cette approche permet de rendre compte des propriétés d'un grand nombre de mouvements. Cependant, une expérience de Konzem (1987) a mis en évidence que cette approche n'était pas suffisante pour expliquer tous les phénomènes. Dans cette expérience réalisable très simplement, aussi appelée expérience Gamma-V, on demande d'abord aux sujets de s'entraîner à dessiner avec leur main non dominante la lettre γ très rapidement. Une fois que la lettre est dessinée avec assez de dextérité, on demande alors aux sujets de dessiner avec la main dominante la lettre v très rapidement. Cette lettre est très proche de γ mais ne possède pas de boucle dans sa partie basse, ce qui implique une planification du mouvement complètement différente. Les sujets n'ont semble-t-il aucun problème pour dessiner ces deux formes séparément. Au regard de l'approche précédente basée sur les invariances, on peut faire l'hypothèse qu'un schéma moteur est mémorisé pour chacune de ces lettres. Si maintenant, on demande aux sujets de dessiner ces deux formes en même temps, i.e. γ avec la main dominante et v avec la main non-dominante, Konzem (1987) a alors observé que les sujets avaient tendance à dessiner la même forme avec leurs deux mains, soit γ soit v .

Cette expérience montre donc que l'approche de la production motrice comme contrôle des invariances ne s'applique pas dans cette situation. En effet, les sujets étant capables de produire chaque lettre séparément, ils devraient être capables de les réaliser de façon synchrone en contrôlant les deux programmes moteurs séparément. Afin de modéliser ce type de phénomène, il est donc nécessaire d'introduire le contexte⁹ dans lequel est réalisé chaque programme moteur, et les contraintes biomécaniques qu'il implique. L'approche dynamique qui va être présentée ici s'est ainsi attachée à rendre compte de ces deux effets, d'une part l'influence du contexte, et d'autre part l'influence des contraintes biomécaniques.

9. Mouvement réalisé par une seule main ou mouvement bimanuel.

Stabilité & auto-organisation des mouvements

L'expérience présentée précédemment a mis en évidence que lorsque l'on produit des mouvements simultanément avec nos deux mains, leurs trajectoires sont liées de façon presque inexorable. Ce résultat montre que pour comprendre ce type de comportement, on ne peut pas considérer les mouvements de deux membres de façon séparée. Le système moteur semble en effet contrôler les deux de façon simultanée pour arriver à produire un mouvement répétitif stable au niveau spatial et temporel.

Pour tenter de modéliser ce type de comportement, des études ont considéré la production des mouvements humains comme un système complexe. Ce type de système met en jeu un très grand nombre de paramètres et l'observation de leur comportement met en évidence l'existence d'états stables qui semblent attirer le système en fonction des différents paramètres. De plus si le système s'écarte de ces états, il y revient de façon inexorable. La possibilité de faire émerger des états stables dans un système contenant un très grand nombre de paramètres nécessite donc que tous les paramètres du modèle s'organisent spontanément en fonction des contraintes pour arriver à rester autour de cet état (Haken, 1983).

Le cas des mouvements et l'exemple présenté précédemment semblent mettre en évidence que les mouvements humains peuvent être modélisés par ce type de comportement. Quand une personne produit un mouvement particulier, les paramètres mécaniques qui contraignent les caractéristiques spatio-temporelles semblent attirés vers un état de stabilité optimale.

Systèmes dynamiques, attracteurs & phase relative

Ce type de comportement est présent dans de nombreux domaines, allant de la physique des instruments de musique, la flûte à bec par exemple (Terrien *et al.*, 2013), à la modélisation du fonctionnement des processus cérébraux en tant que réseaux auto-organisés (Palm et Aertsen, 1986). De façon très générale, ce type de système est appelé système dynamique et peut être défini comme un système dont l'état à un instant t dépend de son état à tous les instants précédents. Dans un cas discret, il peut se définir de la façon suivante :

$$\begin{cases} x(0) = x_0 \\ x(t) = g(x(t-1), x(t-2), \dots, x(t-n)) \end{cases} \quad (1.4)$$

où g est une fonction de n paramètres et x_0 l'état initial du système. Dans le cas des mouvements humains, Haken *et al.* (1985) ont montré que l'on pouvait modéliser les équations du mouvement de la coordination bimanuelle de deux doigts faisant des mouvements répétitifs (voir figure 1.3) par le système suivant :

$$\begin{cases} \ddot{x}_1 + f_1(x_1, \dot{x}_1) = I_{12}(x_1, x_2) \\ \ddot{x}_2 + f_2(x_2, \dot{x}_2) = I_{21}(x_1, x_2) \end{cases} \quad (1.5)$$

où x_1 et x_2 sont les positions de chaque doigt au cours du temps, f_1 et f_2 deux fonctions dépendant des caractéristiques biomécaniques des articulations et des muscles, et I_{12} et I_{21} des fonctions définissant le couplage entre les deux doigts. En dérivant ce système, Haken *et al.* (1985) ont pu ainsi montrer que le mouvement des deux doigts pouvait être modélisé par deux oscillateurs harmoniques :

$$\begin{cases} x_1(t) = r_1 \cos(\omega t + \phi_1) \\ x_2(t) = r_2 \cos(\omega t + \phi_2) \end{cases} \quad (1.6)$$

où r_1 et r_2 correspondent à l'amplitude du mouvement de chaque doigt, ω leur fréquence, et ϕ_1 et ϕ_2 leurs phases. Haken *et al.* (1985) ont ainsi défini une grandeur essentielle, $\phi = \phi_1 - \phi_2$, appelée phase relative. Kelso (1997) a lui montré que le comportement de ce système était attiré par deux phases relatives particulières, $\phi = 0$ degrés ou $\phi = 180$ degrés (i.e. soit les doigts bougent en phase, soit ils bougent en antiphasé).

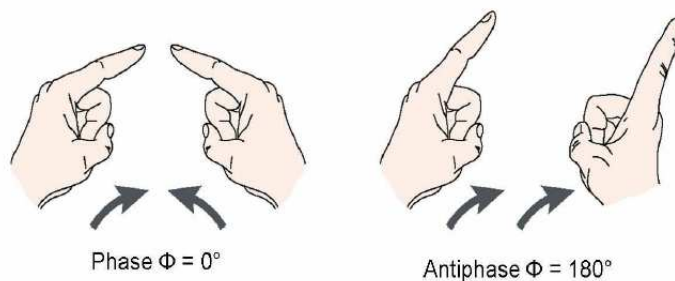


FIGURE 1.3 – Etats stables mis en évidence dans la coordination bimanuelle – figure issue de Kelso (1997)

Sans détailler ces études d'avantage, il est important de retenir que ce type de mouvement répétitif peut être entièrement caractérisé par le couplage entre deux oscillateurs harmoniques et que leurs états stables, aussi appelés attracteurs moteurs, sont caractérisés par la phase relative entre les deux oscillateurs.

Dans le cas des mouvements graphiques, des études ont également mis en évidence l'existence de tels états stables dans la production des mouvements, et en particulier dans la production de mouvements répétitifs par exemple les mouvements elliptiques. Les études sur la production des mouvements elliptiques seront détaillées en chapitre 3 où l'on s'intéressera à la relation entre le son de frottement produit quand une personne dessine une ellipse et l'ellipse qui est réellement imaginée par le sujet. L'idée importante qu'il faut retenir est que la production des mouvements graphiques peut être abordée avec l'approche dynamique qui met en évidence l'importance de la phase relative dans le contrôle de la production des mouvements. De plus, dans le cas des ellipses, cette phase relative ϕ est directement reliée à l'excentricité de la forme dessinée par la formule suivante¹⁰ :

$$\phi = \sqrt{1 - \tan^2 \frac{\phi}{2}}.$$

Un point essentiel à noter également est que ce type de mouvements respecte la loi en puissance $1/3$, caractéristique des mouvements biologiques¹¹. L'approche dynamique ne remet donc pas en cause les invariances présentées précédemment, mais propose juste une explication différente de l'origine de ces invariants et sur le contrôle du mouvement. Selon l'approche dynamique, la loi en puissance $1/3$ est la conséquence de la structure biomécanique des membres mis en jeu. Elle émerge naturellement du couplage entre deux oscillateurs harmoniques et n'est donc pas contrôlée par le système nerveux central à plus haut niveau. Cette approche permet donc de relier directement la géométrie des mouvements à leurs dynamiques.

1.4 Perception des mouvements biologiques

La section précédente a présenté les mécanismes impliqués dans la production des mouvements graphiques. On a vu dans un premier temps que les mouvements biologiques pouvaient être décrits en terme d'invariance cinématique, et en particulier que la vitesse tangentielle du mouvement était

10. Une démonstration détaillée est présentée en annexe 7.1.

11. Une démonstration détaillée est présentée en annexe 7.1.

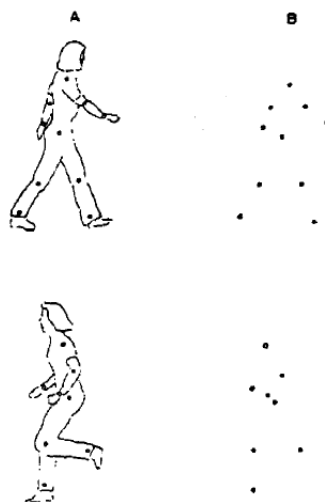


FIGURE 1.4 – Stimuli visuels utilisés dans les expériences de Johansson (1973) sur la reconnaissance visuelle des mouvements biologiques – figure issue de Johansson (1973).

reliée à la courbure de la trajectoire parcourue par la loi en puissance $1/3$. Dans un second temps, on a vu que les mouvements biologiques étaient la conséquence de contraintes biomécaniques et que les invariances cinématiques semblaient être la conséquence de ces contraintes biomécaniques. Les mouvements biologiques ont donc des propriétés très spécifiques. Nous avons présenté ces connaissances afin de mieux comprendre quelles informations peuvent être extraites des sons de frottement produits quand une personne dessine. Du point de vue perceptif, l'approche écologique a elle suggéré l'importance de la motricité dans les caractéristiques des processus perceptifs. Dans cette section, nous allons présenter des résultats issus d'expériences s'intéressant principalement à la perception visuelle des mouvements biologiques. Les invariants mis en évidence précédemment, et en particulier la loi en puissance $1/3$ contraignent en effet de façon très forte notre perception visuelle du mouvement. On pourra par la suite se demander s'ils ne peuvent pas également être extraits des sons de frottement produits par un mouvement graphique.

1.4.1 La loi en puissance $1/3$ contraint la perception visuelle de la vitesse.

La perception des mouvements biologiques est un sujet qui a été étudié depuis très longtemps. On peut mentionner en particulier les travaux de Johansson (1973) qui ont mis en évidence que l'on est capable de reconnaître le mouvement d'une personne juste en regardant des points lumineux positionnés à des endroits bien particuliers de son corps (voir figure 1.4). Le système visuel semble donc particulièrement adapté pour la perception des mouvements biologiques. Par ailleurs, une étude de Runeson (1974) a montré que lorsque l'on observe un point se déplacer à vitesse constante on perçoit des accélérations quand le point effectue des allers retours sur un trait. La perception d'une vitesse constante ne semble donc pas être cohérente avec la notion de vitesse tangentielle constante du point de vue physique.

Une étude qui a complètement changé la façon d'aborder la perception visuelle des mouvements est l'étude réalisée par Viviani et Stucchi (1992) qui ont mis en évidence que quand un point se déplace le long d'une trajectoire, la perception d'une vitesse constante, aussi nommée vitesse uniforme, est contrainte par la loi en puissance $1/3$. Dans leurs expériences, les sujets devaient calibrer la vitesse d'un point lumineux parcourant soit des ellipses de différentes excentricités, soit des formes pseudo-aléatoires (voir figure 1.5). Les sujets calibraient en fait implicitement l'exposant β de la loi en puissance reliant la vitesse tangentielle du mouvement à la courbure de la trajectoire parcourue,

$v_t(s) = KC(s)^{-\beta}$ où s est l'abscisse curviligne. Viviani et Stucchi (1992) ont observé que pour percevoir le mouvement le plus uniforme/constant possible, les sujets avaient calibré l'exposant β de la loi autour de la valeur $1/3$. Cela révèle donc que cet invariant cinématique caractéristique des mouvements biologiques semble également être un invariant perceptif qui contraint la perception visuelle des mouvements.

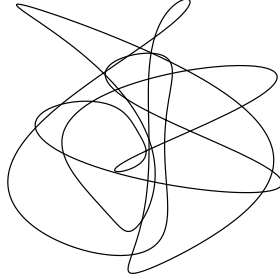


FIGURE 1.5 – Exemple de forme pseudo-aléatoire.

Cette expérience est très intéressante car elle va dans le sens des approches écologiques et éenactives. En effet, le percept correspondant à la notion de vitesse constante ne correspond pas à la réalité physique de vitesse constante mais à la façon dont nous produisons nos mouvements, ce qui peut suggérer que le système visuel a évolué de façon à s'adapter à cette contrainte motrice. de'Sperati et Stucchi (1995) ont de plus vérifié que cette loi n'est pas la conséquence du fonctionnement des mouvements oculaires, suggérant donc que l'origine de cette illusion doit être cherchée ailleurs que dans cette propriété fonctionnelle du système oculomoteur.

Dans l'objectif de déterminer l'origine de cette loi, des études mathématiques ont cherché à modéliser les mouvements dans un autre espace que l'espace euclidien définissant la notion de vitesse tangentielle. En effet, si l'on ne perçoit pas de façon uniforme les mouvements dont la vitesse tangentielle est constante c'est peut-être que l'espace euclidien n'est pas le plus adapté pour représenter les informations cinématiques extraites par le système visuel. Des études très anciennes ont déjà mis en évidence que l'espace euclidien n'était pas nécessairement l'espace mathématique le plus approprié pour modéliser les informations extraites par le système visuel (Luneburg, 1950). De même il ne semble pas être le plus adapté pour représenter les caractéristiques spatio-temporelles des mouvements graphiques (Biess, 2013). Concernant la loi en puissance $1/3$, différentes études ont montré que l'espace affine semblait prédire les observations comportementales faites par Viviani et Stucchi (1992).

Les opérations géométriques considérées dans un espace euclidien sont des translations et des rotations, alors que dans un espace affine, les opérations géométriques considérées sont des dilations et des translations. En ce qui concerne le système visuel, l'espace affine semble plus adapté pour représenter ce qui est perçu. En effet, quand on voit un objet en rotation et en translation, il est ensuite projeté de façon parallèle sur la rétine. La projection sur la rétine implique une dilatation, qui n'est pas directement prise en compte dans la structure mathématique définissant un espace euclidien. De la même façon que l'on peut définir un objet qui se déplace par sa vitesse tangentielle dans un espace euclidien, il est possible de le définir par une *vitesse affine* dans un espace affine. Pollick et Sapiro (1997) ont montré qu'il était possible de relier la vitesse affine v_a à la vitesse tangentielle v_t par la relation suivante :

$$v_a(s) = v_t(s)C(s)^{-1/3} \quad (1.7)$$

où C est la courbure de la forme et s l'abscisse curviligne. Si on considère désormais un mouve-

ment respectant la loi en puissance $1/3$, $v_t(s) = KC(s)^{1/3}$ (avec K le gain de vitesse). En injectant v_t dans l'équation 1.7, on obtient une vitesse affine constante $v_a = K$. La métrique affine semble donc particulièrement adaptée pour refléter la perception de la vitesse par le système visuel.

Dans une perspective écologique, on peut dire que la loi en puissance $1/3$ semble être un invariant capté par le système visuel pour optimiser la perception des mouvements humains. Sans s'avancer sur l'origine de cette adaptation au niveau visuel, encore source de débat, il est important de noter que la métrique affine semble bien adaptée pour décrire la façon dont nous percevons les mouvements au niveau visuel. Nous allons voir dans la section suivante que la loi en puissance $1/3$ contraint également la perception de la géométrie d'une forme parcourue par un point lumineux.

1.4.2 La loi en puissance $1/3$ contraint la perception de la géométrie d'une trajectoire.

Quand on perçoit un point lumineux en mouvement, notre mémoire visuelle devrait nous permettre de déduire la géométrie de la trajectoire à partir du mouvement d'un point lumineux qui parcourt une trajectoire. Viviani et Stucchi (1989) se sont ainsi intéressés à la perception de la géométrie d'un point lumineux parcourant une trajectoire circulaire en fonction de différentes conditions cinématiques. Dans une expérience, les sujets devaient juger si une ellipse parcourue par un point lumineux était verticale ou horizontale. Si la réponse était correcte alors l'excentricité de l'ellipse diminuait, convergeant ainsi vers un cercle, qui en théorie n'a plus d'orientation. L'ellipse était parcourue par un point lumineux se déplaçant selon différentes conditions cinématiques. Dans une condition, le point lumineux se déplaçait à vitesse constante le long de l'ellipse (quelle que soit l'excentricité), dans deux autres conditions, le point lumineux parcourait l'ellipse en accélérant dans les parties opposées de l'ellipse, soit les parties verticales, soit les parties horizontales. L'hypothèse qu'ont voulu tester Viviani et Stucchi (1989) était que la perception de la forme était contrainte par la covariation entre la courbure et la cinématique, et plus spécifiquement que quand le mouvement ne respecte pas la loi en puissance $1/3$, i.e. avec des accélérations dans les parties verticales ou horizontales, le cercle est perçu comme une forme elliptique.

Les résultats ont bien confirmé cette hypothèse et ont en particulier mis en évidence que quand la cinématique du mouvement était non naturelle, le cercle était perçu de façon significativement plus elliptique qu'en réalité (voir figure 1.6). Cette expérience met donc en évidence que la perception de la forme parcourue par un point lumineux est complètement contrainte par la cinématique du mouvement du point lumineux. Ce résultat est très intéressant car il montre que la façon dont on perçoit une géométrie est contrainte à la fois par la géométrie réellement parcourue, mais également par des contraintes cinématiques correspondant à nos caractéristiques motrices, ici la loi en puissance $1/3$. Cette expérience a également été reproduite dans le cas de la modalité haptique. Des sujets devaient juger de l'horizontalité ou de la verticalité de la trajectoire effectuée par leur bras qui était mis en mouvement par un bras articulé reproduisant les mêmes trajectoires et cinématiques que dans l'expérience visuelle de Viviani et Stucchi (1989). Viviani *et al.* (1997) ont alors mis en évidence les mêmes résultats que dans la modalité visuelle. Un biais sur la perception de la circularité de la forme est apparu pour des stimulations non biologiques. De plus une expérience de tracking d'un point visuel en boucle fermée (main visible) a mis en évidence que les sujets étaient plus habiles pour suivre précisément un mouvement aléatoire s'il respectait la loi en puissance $1/3$ (Viviani *et al.*, 1987; Viviani et Mounoud, 1990). Viviani *et al.* (1997) ont alors proposé que les mouvements biologiques sont encodés à un niveau amodal, ne relevant *a priori* pas d'une modalité spécifique. Il est intéressant de noter qu'aucune expérience n'a jusqu'à présent tenté de montrer cette hypothèse du point de vue de la modalité auditive.

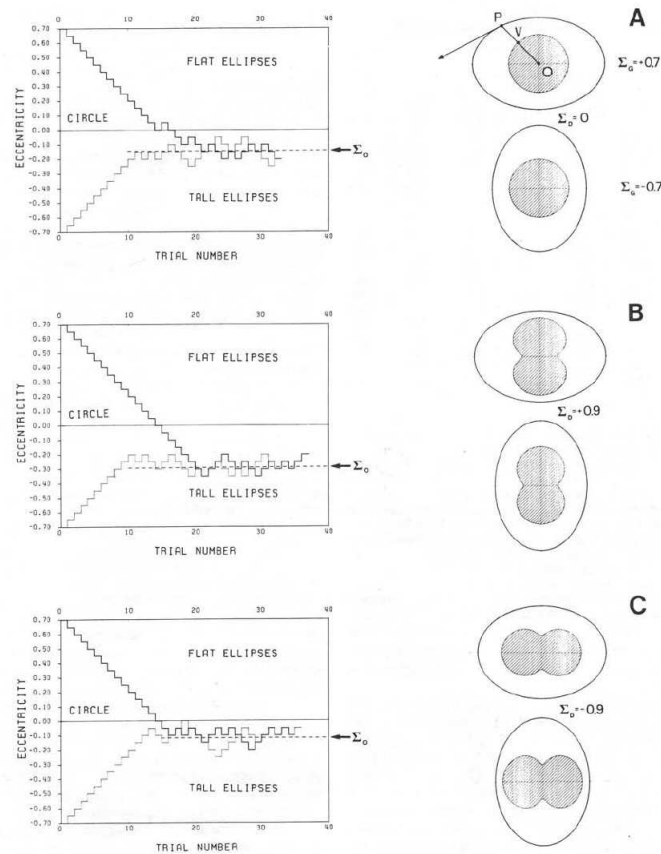


FIGURE 1.6 – Résultats de l'expérience réalisée par Viviani et Stucchi (1989). A gauche on peut voir des résultats typiques pour différents sujets, à droite on peut voir les deux ellipses extrêmes dans les tâches d'ajustements et les profils de vitesse associés (zones grisées). On peut voir sur les panels A, B et C que l'excentricité correspondant au seuil de discrimination, et donc à la perception d'une trajectoire circulaire, est significativement décalée vers une ellipse orientée verticalement pour les conditions cinématiques non naturelles – figure issue de Viviani et Stucchi (1989).

Il est important de noter que les situations qui vont être étudiées dans cette thèse du point de vue de la modalité auditive ne seront pas équivalents à cette situation visuelle, en effet, les variations de timbre d'un son de frottement (monophonique) traduisent uniquement la vitesse d'un mouvement mais ne traduisent en aucun cas des indices spatiaux comme c'est le cas dans la perception du mouvement d'un point lumineux.

1.5 Perception et catégorisation des formes géométriques

Dans la section précédente, on a vu que la perception visuelle d'une forme géométrique *dynamique* (i.e. parcourue par un point lumineux) était contrainte par les caractéristiques des mouvements biologiques. Mais avant d'être caractérisé par des sons de frottement, les mouvements graphiques sont avant tout caractérisés par une trace visuelle. Dans les études présentées dans la thèse, nous nous sommes ainsi intéressés à la perception et l'imagination d'une forme géométrique à partir d'un son de frottement évoquant un mouvement biologique. Au-delà des aspects dynamiques liés au geste, une forme géométrique est avant tout définie par une géométrie *statique*. On peut ainsi se demander si plus que le geste, la forme de la trajectoire parcourue par un geste n'est pas reconnaissable à travers les sons de frottement produits par une personne qui dessine. Pour aborder ce problème, nous allons présenter des études relevant de la perception et la catégorisation visuelle de formes géométriques. Il semble en effet que notre perception visuelle des formes géométrique et la façon dont nous les catégorisons soient reliées d'une part à la façon dont on les produit et d'autre part à une expérience visuelle très forte.

1.5.1 La perception d'une forme géométrique : une action simulée

On a vu dans la section 1.1.3 que les processus de perception semblaient correspondre à une simulation de l'action réalisée. Dans le cas spécifique des formes géométriques cela a été mis en évidence aussi bien au niveau comportemental que neuropsychologique. Freyd (1983a,b) a par exemple mis en évidence que les représentations motrices sont intrinsèquement reliées à la perception d'une forme visuelle. Freyd (1983a) a dans un premier temps demandé à des sujets de mémoriser des photos où l'on voyait une personne en train de sauter, puis dans un second temps les sujets devaient reconnaître la photographie dans une série de photos composées soit de distracteurs où la photographie était prise juste avant la scène mémorisée ou juste après. Elle a alors observé que les sujets étaient meilleurs pour reconnaître la photo mémorisée quand les distracteurs étaient des photos présentant la scène à des instants précédents la scène mémorisée. Cela suggère que pour mémoriser la scène, les sujets ont imaginé ce qu'il se déroule après la photo, c'est-à-dire qu'ils ont en quelque sorte simulé l'action réalisée. Il est ainsi plus difficile de distinguer les photos de la scène visuelle se déroulant après la scène mémorisée que celles se déroulant avant (qui n'ont pas été *a priori* imaginées). Freyd (1983b) a ensuite mis en évidence que l'on était plus habile pour reconnaître une lettre distordue présentée dynamiquement si on avait déjà vu la dynamique qui l'a produite précédemment. Ces deux études suggèrent donc que lorsque l'on mémorise une forme, plutôt que de mémoriser uniquement la géométrie de la forme complète, on mémorise également la dynamique du mouvement réalisé.

Du point de vue neuropsychologique, des études déjà mentionnées dans la section 1.1.3 ont mis en évidence que lorsque l'on perçoit une lettre que l'on sait produire, les aires motrices impliquées lors de la production de la forme sont également activées (Longcamp *et al.*, 2003, 2006, 2008; James et Gauthier, 2006). Cela suggère donc que la mémorisation d'une forme géométrique fait intervenir le système moteur. On peut alors faire le parallèle avec la théorie motrice de la perception de la parole (Liberman et Mattingly, 1985) qui suggère que l'on ne perçoit pas les signaux de paroles comme

des signaux purement acoustiques, mais qu'on les percevrait en terme de gestes phonatoires potentiels. Par ailleurs, une étude récente de Cook *et al.* (2014) a mis en évidence le rôle de la motricité dans la perception visuelle de la phase relative d'un mouvement. Dans une expérience utilisant un paradigme d'interférence, des sujets devaient juger si deux doigts présentés sur un écran se déplaçaient avec le même mouvement. Simultanément, il leur était demandé de taper sur un clavier avec une fréquence particulière correspondant à celle de l'un des deux doigts présentés sur l'écran. Cook *et al.* (2014) ont alors observé que la capacité de discrimination était significativement affectée par cette tâche motrice concurrente. Finalement on peut se demander dans quelle mesure notre système moteur est également impliqué lorsque l'on perçoit et catégorise une forme géométrique statique.

Dans une étude psychophysique récente, Wamain *et al.* (2011) ont mis en évidence que l'on discriminait mieux des ellipses visuelles statiques que l'on dessine naturellement et que ces ellipses correspondent à des ellipses dont l'excentricité est proche de .9, correspondant à une phase relative de 45 degrés si on adopte le modèle proposé par l'approche dynamique (cf. section 1.3.2). Ce résultat est à mettre en parallèle avec les résultats présentés précédemment sur la production des formes elliptiques. Il semble donc que les formes que l'on discrimine le mieux sont des formes qui correspondent aux attracteurs moteurs, ici ceux de l'ellipse.

Dans cette thèse, il sera intéressant de se demander si la perception d'un son de frottement n'implique pas le même processus de simulation d'un mouvement et plus particulièrement si la phase relative entre les oscillateurs harmoniques qui modélisent le mouvement des effecteurs n'est pas également extraite des variations de timbre d'un son de frottement.

1.5.2 Catégorisation des formes géométriques : une expérience visuelle ?

On sait désormais que le système moteur intervient très fortement dans la perception visuelle d'une forme géométrique, et plus généralement dans la perception d'un mouvement. Parallèlement à ces connaissances, des études se sont focalisées plus particulièrement sur la façon dont on catégorisait les formes géométriques. Quand on imagine une forme géométrique, les yeux fermés par exemple, on imagine une forme particulière. De même, lorsque l'on dessine une forme comme un rectangle, on dessine implicitement une forme qui nous semble représenter le mieux cette catégorie de formes géométriques.

Dans des travaux qui font désormais référence sur le sujet, Rosch (1973) a étudié la notion de forme géométrique préférentielle/prototypique chez des individus d'une tribu dont le langage ne possédait aucun mot pour définir le concept de forme géométrique prototypique¹². Elle a observé que les individus de cette tribu apprenaient alors plus rapidement à associer certaines formes géométriques à des mots de leur langage que d'autres. Elle a ainsi fait l'hypothèse que certaines formes étaient plus naturelles que d'autres et a ainsi défini la notion de *forme prototypique*, ou de *catégorie naturelle*, comme étant la forme la plus représentative d'une catégorie de forme particulière. De nombreuses expériences ont ensuite confirmé l'existence de ces formes particulières du point de vue perceptif en utilisant différentes méthodes telles que des jugements de typicité (Rosch, 1975; Feldman, 2000), des expériences d'apprentissage (Rosch, 1975) ou encore de reconnaissance (Theurel *et al.*, 2012). Dans toutes ces expériences, la modalité visuelle a été utilisée pour juger de la prototypicité des formes considérées. Feldman (2000) a par exemple mis en évidence les formes les plus prototypiques de la catégorie des quadrilatères et des triangles (voir figure 1.7). Des expériences ont également utilisé la motricité pour révéler l'existence de ces formes prototypiques. Kalénine *et al.* (2013) ont par exemple montré que des adultes et des enfants dessinaient des rectangles et des triangles

12. Dans la même étude, Rosch (1973) a également étudié la notion de couleur prototypique.

dans des proportions similaires autour d'un prototype particulier. Par exemple pour le rectangle, ils ont montré que le prototype correspondait à un rectangle avec la longueur orientée à l'horizontale et présentant un rapport longueur sur largeur d'environ 2.3.

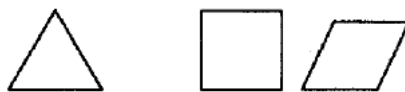


FIGURE 1.7 – Prototypes du triangle et du quadrilatère mis en évidence dans Feldman (2000) – figure issue de Feldman (2000)

Les prototypes géométriques ont également été mis en évidence dans la modalité haptique. Theu-rel *et al.* (2012) ont par exemple montré qu'il existait un prototype haptique du rectangle dans une tâche de reconnaissance de formes effectuée les yeux bandés. Cependant, en reproduisant cette expérience avec des aveugles de naissance, ils n'ont pas pu mettre en évidence l'existence de ce prototype. Cette expérience suggère donc que la modalité visuelle est une condition nécessaire à l'émergence d'un prototype géométrique.

Toutes ces études mettent en évidence que nous catégorisons les formes géométriques autour de formes spécifiques. Ces connaissances sont intéressantes pour les travaux qui vont être présentés dans la thèse. En effet, si nous sommes capables de nous imaginer une forme à partir d'un son, il pourra alors être intéressant de vérifier si la forme que l'on s'imagine à partir du geste évoqué par les sons de frottement correspond également au prototype géométrique de la catégorie considérée. Réciproquement, l'utilisation de la modalité auditive pour évaluer l'existence de prototypes géométriques est intéressante car elle permettra d'évaluer le rôle de la motricité dans l'émergence des prototypes. En effet, les sons de frottement générés quand une personne dessine contiennent principalement des informations dynamiques sur le geste réalisé. On peut alors utiliser ces sons pour évaluer si la dynamique d'un geste est prototypique en soi. Et dans l'éventualité où l'on mettrait en évidence l'existence d'un prototype, on pourra alors évaluer si ce prototype est commun avec un prototype visuel ou moteur.

1.6 Contexte d'étude : analyse par la synthèse de sons de frottement

Sur la base des points de vue décrits précédemment aussi bien du point de vue de la perception sonore, et plus particulièrement de l'approche écologique, que des sciences du mouvement, et plus particulièrement sur la production et perception des mouvements biologiques, nous proposons d'aborder les relations entre les gestes et les sons par le biais de la synthèse sonore de sons de frottement qui est une méthode permettant de contrôler les ressentis propres aux gestes du point de vue de l'audition. L'utilisation de la synthèse sonore est ainsi le point central de tous les travaux présentés dans cette thèse. Comme cela a été évoqué en introduction ce sera à la fois notre outil d'étude mais également le centre de nos questionnements et pour aborder ces problèmes nous allons utiliser une méthode bien connue appelée analyse par la synthèse.

La synthèse sonore permet de créer des sons de frottement en contrôlant précisément la cinématique du mouvement qui a généré le son (nous allons présenter le modèle dans la suite de cette section). L'intérêt de la synthèse est ainsi de permettre de contrôler ces informations *a priori* et de pouvoir les modifier en temps réel. Ensuite, par le biais de différents tests perceptifs, il est possible de déterminer quelles sont les informations les plus pertinentes du point de vue perceptif, i.e. les invariants. Cette méthode est souvent appelée analyse par la synthèse (Risset et Wessel, 1999) et a déjà été utilisée pour déterminer les attributs perceptifs du timbre qui permettent la reconnaissance

des instruments de musique (Risset, 1965; McAdams et Cunible, 1992; McAdams *et al.*, 1999). Risset (1965) s'est intéressé par exemple à la perception des sons produits par les instruments de la famille des cuivres comme la trompette. En utilisant la synthèse additive¹³, il a ainsi pu montrer l'importance de l'attaque du son dans la reconnaissance de ce type d'instrument. Afin d'utiliser cette méthode, nous allons présenter un modèle de synthèse de sons de frottement qui permettra d'évaluer l'importance du profil de vitesse dans la perception des sons produits par une personne qui dessine. Pour établir ce modèle, nous allons dans un premier temps présenter les rudiments de physique du frottement nécessaires.

1.6.1 Acoustique du frottement

Les bruits générés par frottement ont été largement étudiés et modélisés car ils sont omniprésents dans notre quotidien. Que ce soit la gêne occasionnée par les bruits de freinage de véhicules, le crissement d'une craie sur un tableau ou plus généralement le bruit d'une surface grattée par un autre objet, tous ces sons nous informent sur l'état, la nature, ou encore la dangerosité d'un objet et d'une action. Dans le but de synthétiser ces sons, il est nécessaire de comprendre les mécanismes qui en sont les causes.

Qu'est-ce qu'un frottement ?

L'étude du contact frottant est présente dans de nombreux domaines de la physique, elle est liée à une multitude de problématiques. Que ce soit pour l'optimisation d'un mécanisme de freinage ou pour l'étude des frottements d'un archet sur une corde de violon, la compréhension et la modélisation des frottements entre deux corps est essentielle. Ces différentes problématiques exigent une précision plus ou moins fine. Les types de frottements que l'on souhaite intégrer dans les modèles de synthèses peuvent également être modélisés par des lois phénoménologiques¹⁴ simples ne nécessitant pas une étude physique précise d'éventuelles non linéarités. Léonard De Vinci, au début du XVIe siècle, Amontons en 1699 et Coulomb (1821) ont mis au point des modèles empiriques simples toujours utilisés pour modéliser les frottements entre deux corps.

Ces modélisations dérivent de la formulation du contact unilatéral entre deux corps qui stipule que, si deux corps entrent en contact, il n'y a pas d'interpénétrabilité entre eux (i.e. l'un des deux corps ne peut pas rentrer dans l'autre). La réaction \vec{R} d'un corps vis-à-vis de l'autre peut alors se décomposer en une composante tangentielle \vec{F}_T et une composante normale \vec{F}_N . Les lois de frottement établies par Coulomb et Amontons établissent une relation entre \vec{F}_N et \vec{F}_T et définissent un coefficient μ , appelé coefficient de frottement, caractérisant la limite entre l'adhérence et le glissement.

$$\begin{cases} \text{si il y a adhérence,} & \|\vec{F}_N\| \leq \mu \|\vec{F}_T\| \text{ et } v_r = 0 \\ \text{si il y a glissement,} & \|\vec{F}_N\| \geq \mu \|\vec{F}_T\| \text{ et } v_r = -K \|\vec{F}_T\| \end{cases}$$

où v_r est la vitesse relative entre les deux corps et K un nombre réel positif.

Des lois dérivées font intervenir un coefficient μ variable selon la vitesse relative des deux corps (Canudas de Wit *et al.*, 1995). En effet, aux faibles vitesses, l'adhérence entre les deux corps sera plus forte, le coefficient μ sera donc plus grand. Euler fait intervenir un coefficient de frottement statique μ_s afin de distinguer la phase de perte d'adhérence de celle de frottement. Cette formulation est

13. La synthèse additive est une technique qui permet de créer des sons en sommant ses différentes composantes harmoniques et en contrôlant l'amplitude de chacune d'entre elles : $s(t) = \sum_n A_n \sin(\omega_n t)$.

14. Le terme phénoménologique ne fait pas référence ici à l'approche phénoménologique de la perception mais à la façon d'aborder un problème physique en considérant des observations empiriques des phénomènes plus que des considérations physiques très précises.

encore couramment utilisée en mécanique car elle ne fait intervenir qu'un seul paramètre μ , ce qui permet une utilisation aisée dans des cas concrets.

État de surface et paramètres cinématiques. Une modélisation très générale des interactions entre deux corps, issue d'une étude tribologique¹⁵, fait état des échanges énergétiques qui ont lieu entre les deux corps en contact (Linck, 2005). Ce sont ces échanges qui peuvent être à la source des bruits créés lors du frottement.

Les échanges prépondérants entre les deux corps sont des transferts d'énergie thermique et d'énergie potentielle (Akay, 2002). Ils sont également liés aux déplacements et à l'intensité du contact entre les deux corps, mais aussi à l'état de surface des deux matériaux.

- **L'état de surface** des deux matériaux en contact est défini par plusieurs paramètres caractérisant la géométrie et l'aspect de la surface. La norme ISO-25178 (2010) précise l'analyse de l'état d'une surface. Dans le cadre de notre étude, une seule grandeur, la rugosité, sera utilisée. Selon les domaines, plusieurs définitions de ce paramètre existent. Pour notre étude, une définition adaptée est :

Rugosité : ensemble des irrégularités d'une surface à caractères micrographiques et macrographiques.

La rugosité est donc définie par des distances moyennes caractérisant différentes échelles de la géométrie (Ren *et al.*, 2010). Il est possible de définir une surface efficace de contact entre les deux corps qui va déterminer à quel *niveau* la rugosité va être *sondée*, et donc dans quelle mesure et comment les corps vont être excités.

- **Les déplacements** des deux corps en contact peuvent être entièrement décrits par les paramètres cinématiques que sont la position, la vitesse, et l'accélération. Généralement, on s'intéresse au frottement d'un corps mobile sur un corps fixe, un seul jeu de paramètres est alors nécessaire et la vitesse considérée est la vitesse relative v_r du corps en mouvement vis-à-vis du corps immobile.
- **L'intensité** du contact entre les deux corps joue un rôle important dans les transferts d'énergie et sur le type de bruit généré. Dans une étude sur différents types de friction, Akay (2002) présente le cas d'une tige cylindrique appuyée sur un disque tournant. Selon l'intensité de la force de contact entre la tige et le disque et selon la vitesse de rotation du disque, différents types de bruits sont émis. Par ailleurs, il est clair que plus la force d'interaction est importante, plus les échanges thermiques et visco-élastiques entre les deux corps seront importants.

Le frottement d'un objet en mouvement, désormais appelé exciteur, sur un objet fixe, désormais appelé résonateur, peut donc être entièrement caractérisé si l'on connaît ses paramètres cinématiques et la rugosité de la surface. Ce frottement peut générer des bruits large bande, générés au niveau du contact entre la surface rugueuse et l'extrémité de l'exciteur et excitant les modes propres du résonateur.

15. La tribologie est la science qui étudie les phénomènes susceptibles de se produire entre deux systèmes matériels en contact, immobiles ou animés de mouvements relatifs.

Modèle phénoménologique

Le bruit large bande généré lors du contact frotté entre un excitateur et une surface rugueuse peut être assimilé à une suite d'impacts entre la tige et toutes les aspérités de la surface. (Ren *et al.*, 2010; Avanzini *et al.*, 2005; Van Den Doel *et al.*, 2001). A chaque passage de l'excitateur au niveau d'une aspérité, un impact a lieu et un son est émis. Le bruit large bande est principalement la conséquence de deux phénomènes :

- l'interaction entre l'excitateur et les aspérités ;
- les modes propres de la plaque qui sont excités à chaque passage de l'excitateur au niveau d'une aspérité.

Ben Abdelounis *et al.* (2010); Stoimenov *et al.* (2007) ont observé que le niveau de pression¹⁶ et la fréquence centrale des sons générés lors d'un frottement étaient linéairement liés à la rugosité de la surface et à la vitesse de déplacement de l'excitateur. Les platines vinyles donnent une bonne illustration de ce phénomène. En effet, selon la vitesse avec laquelle tourne le disque vinyle, la hauteur du son est modifiée¹⁷ (Van Den Doel *et al.*, 2001).

La composante large bande est présente dans quasiment tous les phénomènes de frottement. Elle est plus ou moins prédominante selon la rugosité de la surface frottée. Par exemple, pour une plaque en métal présentant des stries périodiques rapprochées frottées par une tige en fer, cette composante sera très importante dans le mécanisme de génération du bruit car l'interaction même entre la tige et les stries sera très forte.

Le bruit généré par la friction de la mine sur une feuille de papier correspond donc à ce type de bruit de frottement large bande. Les paramètres dynamiques qui influencent le timbre sont donc principalement au nombre de trois : la vitesse tangentielle du stylo, la pression à laquelle le stylo appuie sur la feuille et enfin la rugosité de la feuille. On peut voir sur la figure 1.8 la forme temporelle et la transformée de Gabor d'un enregistrement d'un son de frottement d'un stylo sur une table, le profil de vitesse du stylo qui a produit le son est également superposé. On remarque ainsi clairement la relation entre la vitesse du stylo et les variations spectro-temporelles du son.

1.6.2 Synthèse et contrôle perceptif des sons de frottement

Afin de synthétiser des sons de frottement, nous allons utiliser le modèle phénoménologique présenté dans le paragraphe précédent. Pour contrôler intuitivement ce modèle, nous nous appuyons sur l'approche écologique qui stipule l'existence d'invariants perceptifs structuraux reliés à l'objet et d'invariants transformationnels reliés à l'action. Cette section va présenter le modèle qui va être utilisé dans toute la suite de la thèse, ce modèle permet d'intégrer toutes les connaissances théoriques qui ont été présentées dans les sections précédentes. En effet, sa structure répond aux contraintes de l'approche écologique tout en intégrant la possibilité d'être contrôlé par des cinématiques caractéristiques des mouvements biologiques.

Le paradigme action → objet

L'approche écologique de la perception a donc proposé que pour reconnaître un événement sonore nous nous basons principalement sur des informations reliées à l'objet et à l'action. On peut même décrire un son du quotidien, comme le résultat d'une action sur un objet (e.g. frapper une

16. $L_p(dB) = 20 \log_{10} \left(\frac{p_{mes}}{p_{ref}} \right)$, $p_{ref} = 2.10^{-5}$ Pa.

17. En diminuant la vitesse de lecture, le son sera plus grave ; en l'augmentant il sera plus aigu.

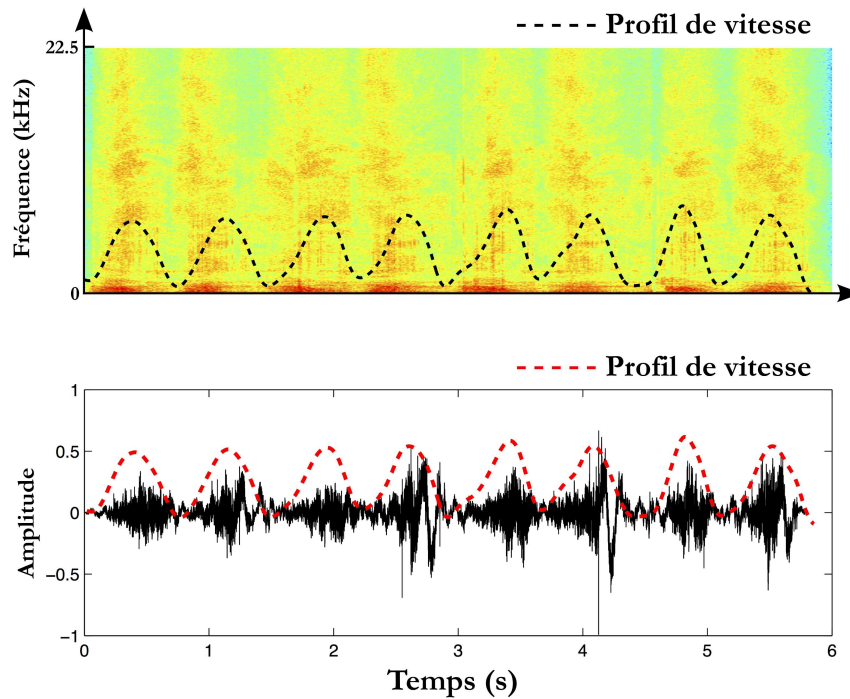


FIGURE 1.8 – Cette figure présente la transformée de Gabor et l'amplitude d'un enregistrement de son de frottement avec le profil de vitesse (traits pointillés) superposé. On voit clairement que le barycentre spectral et l'enveloppe d'amplitude sont liés aux variations du profil de vitesse.

plaque, frotter une vitre...). Dans le cas d'un son de frottement produit par une personne qui dessine, l'action correspond au mouvement du stylo sur les rugosités de la surface et l'objet correspond aux résonances et surtout de la table en dessous de la feuille excitée par les impacts successifs du stylo sur les rugosités de la feuille. Dans l'optique de synthétiser des sons, ce paradigme est très intéressant car il nous fournit tous les éléments de base que notre modèle doit intégrer pour générer un son. Ici, il nous faut donc pouvoir synthétiser l'effet de l'action des impacts de la mine de stylo sur les rugosités et résonances de l'objet. Ce paradigme met donc en évidence l'existence d'un excitateur (la mine du stylo) et d'un résonateur (la feuille et la table) et nous incite naturellement à utiliser un modèle de synthèse source-filtre, particulièrement bien adapté pour notre paradigme.

Un modèle source/filtre

Le paradigme action \rightarrow objet considère donc que le son est le résultat d'une action sur un objet. Du point de vue acoustique, le son $s(t)$ généré par ce type de paradigme peut être vu comme le résultat d'une convolution entre une excitation $e(t)$ correspondant à l'action, par une réponse impulsionnelle $h(t)$, correspondant à la résonance de l'objet : $s(t) = (e * h)(t)$. Dans le cas des sons d'interaction continues comme les frottements, des méthodes d'analyse/synthèse ont mis en évidence que le signal d'excitation $e(t)$ correspondait à un bruit dont la composition spectro temporelle varie en fonction de la vitesse du mouvement, et que $h(t)$ correspondait à la réponse impulsionnelle de l'objet excité (Lagrange *et al.*, 2010). Du point de vue du signal et d'une implémentation en temps-réel, un modèle source-filtre peut être implémenté en filtrant une excitation $e(t)$ par un filtre à réponse impulsionnelle finie dont la réponse impulsionnelle correspond à la réponse réelle $h(t)$ de l'objet. Différentes méthodes existent pour décomposer une réponse impulsionnelle comme un banc de filtres à réponses impulsionnelles finies, on peut en particulier mentionner la décomposition en somme de fractions rationnelles.

Dans notre étude nous nous intéressons au cas spécifique du son produit par le frottement d'une

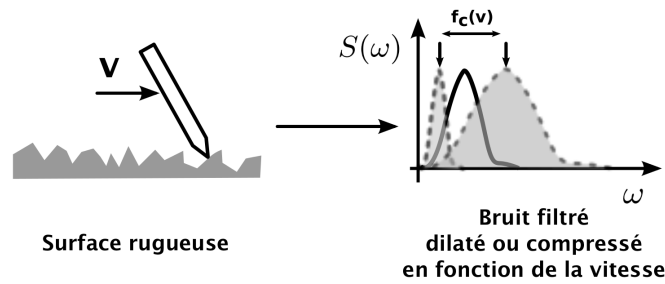


FIGURE 1.9 – La surface frottée peut être représentée comme un bruit filtré passe bas ou passe bande en fonction de ses caractéristiques physiques (principalement sa rugosité). Frotter cette surface à différentes vitesses est équivalent à dilater ou compresser cette bande de bruit en fonction de la vitesse du geste. Cette figure présente un exemple où une surface est représentée par un bruit filtré passe bande dont l’enveloppe spectrale est schématisée sur la droite de la figure (trait plein) et frottée à différentes vitesses (traits pointillés). On voit ainsi dans cette situation que le centre de gravité spectral évolue en fonction de la vitesse. Dans le cas d’un bruit filtré passe-bas, les dilations et compressions feraient varier principalement le centre de gravité spectral mais également l’étalement spectral en fonction de la vitesse.

mine de stylo sur une feuille posée sur une table. Dans ce cas, la source, correspondant à l’interaction entre la mine et les aspérités de la surface, peut être modélisée par un bruit filtré passe-bas (ou passe-bande) dont la fréquence de coupure (ou centrale) est liée linéairement à la vitesse tangentielle du geste (voir figure 1.9). Le filtre, correspondant à la résonance de la table sous la feuille, peut être modélisé par un banc de filtres résonnants dont les fréquences centrales et les amortissements sont calibrés en fonction des modes de résonance de la table. Ainsi, on peut synthétiser très simplement un son de frottement juste en connaissant la vitesse tangentielle du mouvement et en connaissant les résonances de la feuille et de la table. La modélisation et le contrôle de ce banc de filtres ont été étudiés au Laboratoire de Mécanique et d’Acoustique dans le cadre d’études menées sur la perception auditive des sons d’impacts. Une correspondance entre le matériau perçu et les valeurs des fréquences et des amortissements des filtres résonnants a pu être établie et permet ainsi le contrôle intuitif du matériau. Ces travaux ont déjà été décrits précédemment dans la section 1.2. Dans cette thèse, nous nous intéresserons plus particulièrement à la source correspondant à l’action réalisée sur l’objet, le mouvement du stylo dans le cas des sons de frottement. La génération de cette source va être présentée dans le paragraphe suivant.

Description du signal source

Le signal source correspond à un bruit qui, du point de vue phénoménologique, représente les aspérités de la surface que rencontre la mine du stylo lors du frottement. Des études ont mis en évidence que l’on pouvait modifier la rugosité de la surface évoquée en modifiant la composition spectro-temporelle du bruit. Par exemple, les bruits fractaux¹⁸ modélisaient bien les différentes rugosité de la surface. Par ailleurs des études récentes ont mis en évidence que l’on pouvait même contrôler la nature de l’interaction, par exemple est-ce que l’objet frotte, roule ou gratte, et cela en modifiant la structure interne du bruit de façon plus subtile.

En effet, Conan *et al.* (2014a) ont montré qu’en contrôlant judicieusement la corrélation entre l’amplitude et l’espacement temporel entre les impacts simulant les aspérités de la surface du résonateur on pouvait évoquer différents types d’interaction¹⁹. Nous ne présenterons pas ces études et

18. Dont l’amplitude du spectre $S(\omega)$ varie de façon inversement proportionnelle à la fréquence : $S(\omega) = \frac{1}{\omega^\beta}$.

19. Ces résultats sont très intéressants car ils montrent pour la première fois l’importance de la structure interne du bruit dans la perception des interactions continues. Ils ont donc des implications aussi bien dans la modélisation des

nous référons ici aux travaux de thèse de Simon Conan (2014). Dans les études présentées dans la suite, on utilisera des bruits blancs évoquant des bruits de frottements, la statistique des impacts sera donc gaussienne et aucune corrélation temporelle ne sera utilisée pour évoquer des interactions plus spécifiques comme les roulements.

Contrôles perceptifs : mapping entre les paramètres haut-niveau et les paramètres du modèle de synthèse

Afin de relier les paramètres de contrôle haut-niveau comme le type d'interaction (voir figure 1.10), la cinématique du geste ou encore les paramètres de l'objet, aux paramètres bas-niveau du modèle de synthèse, il est nécessaire de faire un mapping entre les différentes valeurs de ces paramètres aux différents contrôles sémantiques (voir figure 1.10). Pour le type d'interaction, dans le cadre d'une collaboration avec Simon Conan *et al.* (2012, 2013, 2014b), nous avons proposé une stratégie de contrôle intuitive basée sur une description sémantique de l'interaction : frottement, grattement, ou roulement. Cette stratégie permet en particulier de réaliser des morphings perceptifs continûment entre ces différentes interactions en modifiant la statistique des impacts (i.e. amplitudes et intervalles temporels entre les impacts). Dans le cas de l'objet, un mapping entre les valeurs des fréquences et des amortissements du banc de filtres résonnants et les descripteurs sémantiques reliés au matériau, à la géométrie et à la taille de l'objet a été réalisé par Aramaki *et al.* (2011).

Dans le cas de la cinématique du geste, on a vu que pour contrôler la vitesse de la mine du stylo, le bruit source est filtré soit par un filtre passe-bas soit passe-bande dont la fréquence de coupure ou centrale est reliée à la vitesse tangentielle de l'objet frottant. Le mapping entre la vitesse tangentielle du mouvement et la fréquence de coupure (ou centrale) f_c du filtre passe-bas (-bande) est définie par un coefficient α reliant linéairement ces deux grandeurs au cours du temps : $f_c(t) = \alpha v_t(t)$. Du point de vue acoustique, augmenter le mapping provoque une augmentation du barycentre spectral du son de frottement. Par ailleurs, ce coefficient rend compte de deux phénomènes physiques : d'une part la rugosité de la surface, mais également de la vitesse relative entre les deux objets en contact, ici entre la mine du stylo et les aspérités de la feuille de papier. Du point de vue perceptif, une modification de ce paramètre peut donc être perçue de deux façons, soit comme une modification de la vitesse moyenne, soit comme une modification de la rugosité de la surface. Dans les expériences présentées dans cette thèse, le mapping sera fixé arbitrairement.

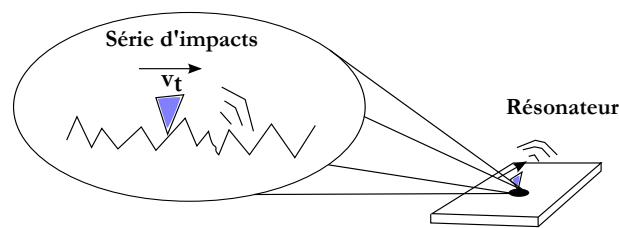
Finalement, cette modélisation du frottement permet de ne contrôler la synthèse du son en ne manipulant que la vitesse tangentielle de la mine du stylo. Cela est très intéressant pour générer des stimuli parfaitement contrôlés dans l'optique d'évaluer l'influence de la cinématique sur la perception des sons de frottement. Nous avons désormais présenté toutes les caractéristiques du modèle nécessaire pour synthétiser des sons de frottement. Ce modèle se base donc sur un paradigme perceptif, le paradigme action \rightarrow objet, proche de la façon dont nous analysons des sons de frottement. Finalement, pour pouvoir complètement aborder le problème des relations entre sons et gestes avec ce modèle, il ne reste plus qu'à le contrôler avec des cinématiques correspondant à des cinématiques de geste.

Modèle de geste

C'est donc ici qu'interviennent les connaissances relevant des sciences du mouvement nécessaires pour l'investigation du geste perçu à travers un son de frottement. Le modèle de synthèse que l'on

phénomènes physiques de roulement, que dans la compréhension de la façon dont le système perceptif extrait les statistiques temporelles d'un signal. Ce dernier point étant un sujet très étudié actuellement (McDermott et Simoncelli, 2011; McDermott *et al.*, 2013; Nelken et de Cheveigné, 2013).

(A) Modélisation phénoménologique du frottement



(B) Synthèse et contrôle de sons d'interactions continues

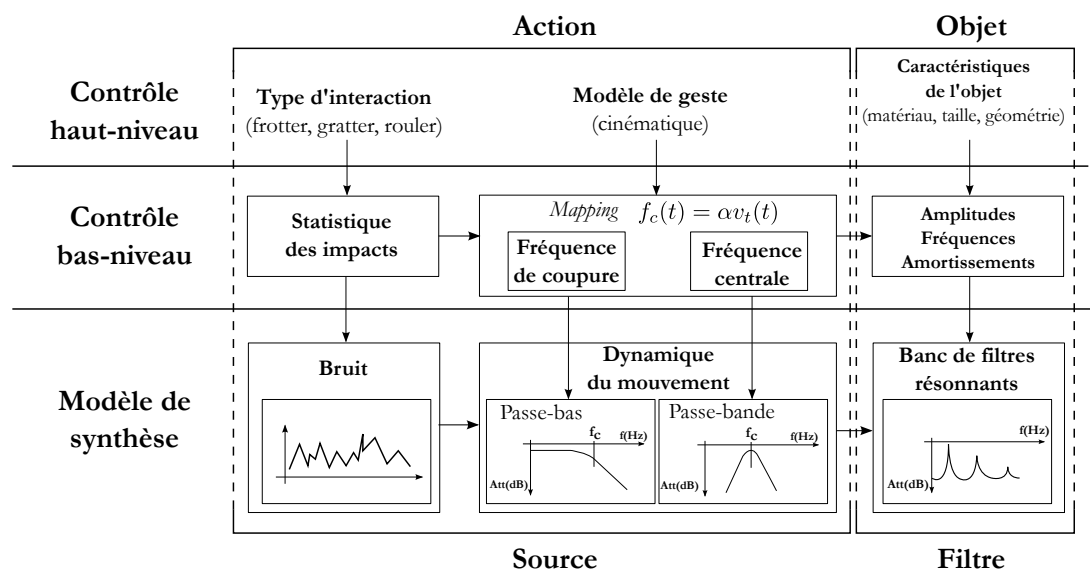


FIGURE 1.10 – Modélisation phénoménologique (Panel A) et modèle de synthèse de sons de frottement s'appuyant sur le paradigme action → objet (Panel B) - Ce schéma présente clairement les différents niveaux de contrôle et d'implémentation du modèle de synthèse. Il est notable que la structure même de ce modèle intègre la façon dont nous percevons les sons et de plus permet le contrôle par des cinématiques biologiques de façon très simple.

vient de présenter permet en effet de générer des sons de frottement en contrôlant tout d'abord les propriétés de l'objet (matériau et taille), en contrôlant la nature de l'interaction (frotter, gratter ou rouler) mais surtout en contrôlant la cinématique d'un geste. Dans cette thèse deux modèles seront principalement utilisés pour générer des profils de vitesse correspondant à des mouvements biologiques.

Tout d'abord en considérant la loi en puissance $1/3$ (présentée en section 1.3.1), et plus généralement la relation de puissance β entre vitesse et courbure $v_t(s) = KC(s)^{-\beta}$ (où s est l'abscisse curviligne, et K le gain de vitesse), il est possible de générer un profil de vitesse à partir d'une forme donnée. A partir d'une forme, on peut donc calculer sa courbure, puis, en définissant la valeur de l'exposant on peut générer des sons de frottement correspondant à une cinématique donnée (correspondant à un mouvement biologique dans le cas où $\beta = 1/3$). Ce modèle de geste sera utilisé dans le chapitre 2 pour étudier la reconnaissance auditive de la loi en puissance $1/3$.

Le second modèle de geste utilisé a été présenté dans la section 1.3.2 et correspond à la modélisation des mouvements graphiques en terme d'oscillateurs harmoniques couplés. Ce modèle est particulièrement intéressant pour générer des mouvements elliptiques et contrôler la géométrie du mouvement par un seul paramètre, la phase relative entre les deux oscillateurs. Quand la phase est nulle, la géométrie correspond à un trait et, quand la phase est égale à 90 degrés, le mouvement est circulaire, toutes les phases relatives comprises entre 0 et 90 degrés correspondent à des ellipses plus ou moins aplaties. Du point de vue cinématique, ce modèle est également très intéressant car le mouvement respecte la loi en puissance $1/3$ (une démonstration est proposée en annexe 7.1). Ce modèle sera utilisé dans le chapitre 3 pour évaluer l'existence de prototype géométrique sonore correspondant à des formes elliptiques. Il sera également utilisé dans le chapitre 4 pour générer des mouvements visuels et des sons de frottement correspondant à des mouvements circulaires ou elliptiques.

Synthèse

Cette section a donc présenté le paradigme, et plus généralement le contexte, d'étude des relations entre gestes et sons qui sera utilisé dans cette thèse. Ce paradigme d'étude, basé sur l'analyse par la synthèse, met ainsi en évidence l'aspect pluridisciplinaire de ces travaux en intégrant à la fois des connaissances perceptives et des connaissances sur la production des mouvements biologiques dans le contexte de la création des sons. Les notions principales à retenir sont les suivantes :

1. le paradigme action \rightarrow objet permet de synthétiser des sons de frottement en découplant les invariants transformationnels reliés au geste des invariants structuraux reliés à l'objet et permet ainsi de générer de stimuli appauvris permettant de comprendre quelles sont les informations pertinentes dans la perception d'un son de frottement ;
2. la flexibilité des modèles source-filtre permet également d'envisager une utilisation en temps-réel des ces outils de synthèse et ainsi de manipuler les caractéristiques du son par des informations dynamiques variant au cours du temps.

1.7 Enjeux

Ce premier chapitre a donc présenté une revue des connaissances nécessaires pour aborder le problème de la perception et de l'interaction avec des sons produits par les mouvements biologiques. Pour conclure cet état de l'art, nous allons présenter les différents enjeux que cela soulève et les problématiques qui seront abordées dans cette thèse.

1.7.1 Enjeux théoriques

Invariant transformationnel associé au geste humain

L'acoustique nous informe donc que les sons produits quand une personne dessine contiennent des informations reliées aux caractéristiques dynamiques des gestes. Par ailleurs, les sciences du mouvement ont pu mettre au jour que les gestes humains respectent certaines règles bien spécifiques, par exemple la loi en puissance $1/3$ reliant la vitesse tangentielle du mouvement à la courbure de la forme dessinée. Du point de vue perceptif, les connaissances sur la perception visuelle du mouvement nous informent que cet invariant cinématique semble être intégré à très haut niveau dans les processus perceptifs. Viviani *et al.* (1997) ont même suggéré que cette loi était encodée à un niveau amodal. On peut donc légitimement se demander si le système auditif permet également de reconnaître cette relation et si ce type de profil de vitesse définit un invariant transformationnel relié à la perception auditive des mouvements humains. Notre paradigme d'étude nous permet donc d'aborder cette question de façon pertinente en contrôlant le ressenti acoustique par des cinématiques spécifiques. Dans le chapitre 2, nous chercherons à évaluer dans quelle mesure cet invariant peut être reconnu par la perception auditive à travers les variations de timbre d'un son de frottement. Il est important de noter la différence avec les expériences de Viviani et Stucchi (1992) qui mettent en jeu des indices visuels spatiaux qui ne seront pas évalués dans nos expériences sonores. Par ailleurs, dans la mesure où la loi en puissance $1/3$ relie les caractéristiques cinématiques du mouvement à la courbure de la trajectoire parcourue, on pourra également se demander dans quelle mesure nous sommes capables d'associer les sons produits par une personne qui dessine à la forme qu'il a vraiment dessinée. Dans le chapitre 2, nous nous poserons dans un premier temps cette question dans une tâche de discrimination de formes lors de laquelle des sujets devront essayer de discriminer des formes géométriques sur la base des sons produits lors du dessin. Le fait de pouvoir re-synthétiser des sons à partir de profils de vitesse particulier nous permettra également de comparer des résultats d'expériences réalisées avec des sons enregistrés avec des résultats des mêmes expériences réalisées avec des sons synthétisés uniquement à partir des profils de vitesse. Nous pourrions ainsi évaluer dans quelle mesure cette information est suffisante pour l'évocation sonore d'un geste.

Révéler les prototypes géométriques par le biais d'une dynamique sonore

Si du point de vue sonore une forme peut donc se définir par un son, cette forme est avant tout caractérisée par une géométrie. Nous avons vu que du point de vue visuel, nous catégorisons et imaginons les formes géométriques autour de formes prototypiques les plus représentatives d'une catégorie particulière. Dans le chapitre 3, nous nous intéresserons ainsi plus spécifiquement à une forme, l'ellipse, et nous essaierons d'évaluer dans quelle mesure un prototype sonore, caractérisé par la dynamique du mouvement, existe. Nous aborderons ce problème en utilisant le modèle dynamique décrit précédemment et en particulier le fait que l'excentricité d'une ellipse est directement reliée à la phase relative entre les effecteurs. Nous essaierons ainsi de déterminer si les attracteurs moteurs dans le cas des mouvements elliptiques ne sont pas également caractéristiques de la géométrie des prototypes, et en particulier du point de vue de la perception sonore. Finalement, cela pose la question de l'origine des prototypes et plus particulièrement de l'influence des contraintes motrices dans l'émergence de la géométrie de ces prototypes. L'utilisation du son et plus particulièrement de la synthèse sonore est un moyen unique d'aborder cette question initialement posée du point de vue psychologique. Elle permet en effet d'évaluer l'influence de la dynamique gestuelle indépendamment des informations géométriques implicitement contenues dans un stimulus visuel. Du point de vue de la synthèse, cette question est également essentielle car elle permet de mieux comprendre

quelle est réellement la forme qu'une personne s'imagine quand elle entend les sons de frottement produits par une personne.

Influence du son sur la perception visuelle et la reproduction motrice des mouvements biologiques

La littérature concernant la perception visuelle des mouvements biologiques a mis en évidence deux illusions visuelles dans lesquelles la vitesse et la géométrie perçues d'un mouvement sont contraintes par la loi en puissance $1/3$. La perception visuelle semble donc s'être adaptée pour percevoir de la façon la plus optimale les mouvements biologiques en *lissant* les variations de vitesse et en intégrant la géométrie à partir des covariations entre la vitesse tangentielle et la courbure. Dit autrement, du point de vue perceptif, une variation de vitesse tangentielle est nécessairement associée à une variation de la courbure de la forme perçue. Du point de vue sonore, les sons de frottement semblent eux transmettre la vitesse tangentielle d'un mouvement de façon très saillante. Il semble donc intéressant d'évaluer dans quelle mesure une variation de vitesse sonore ne peut pas interférer avec la perception visuelle de la vitesse et de la géométrie d'une trajectoire. Dans le chapitre 4, trois expériences aborderont ce problème dans des tâches de synchronisation sensori-motrice. Elles évalueront ainsi dans quelle mesure l'ajout d'un son de frottement évoquant un mouvement sur un stimulus visuel décrivant différentes trajectoires (circulaires & elliptiques) parcourues par un point lumineux avec différentes dynamiques peut modifier la géométrie de la trajectoire visuelle perçue et reproduite. Cette expérience est intéressante aussi bien du point de vue fondamental, car elle permet de mieux comprendre comment on intègre un mouvement audiovisuel, mais également du point de vue appliqué car elle ouvre la possibilité de créer des interfaces multimodales pour le guidage et l'apprentissage basés sur ce paradigme de synthèse sonore.

1.7.2 Enjeux applicatifs

En plus des enjeux théoriques que l'on vient de présenter, différentes applications peuvent être envisagées. Une première application possible concerne le design sonore et la création de sons de façon intuitive. En effet, si nous arrivons à comprendre quels sont les paramètres du geste imaginés à partir d'un son de frottement, nous pourrions proposer des contrôles simples pour la génération des sons de frottement évoquant des mouvements humains. Ces applications potentielles sont à mettre en perspective avec les travaux réalisés par Adrien Merer (2011) dans le cadre du développement de synthétiseurs de sons évoquant des mouvements avec des contrôles basés sur une description graphique et sémantique du mouvement.

D'autre part, une application envisageable concerne l'utilisation du son, et plus particulièrement des sons évoquant des mouvements, pour apprendre et guider des gestes. Cette problématique sera abordée dans un cas particulier de geste, le geste d'écriture. Si on arrive à évoquer des gestes graphiques à partir des sons, pourquoi ne pas les utiliser pour guider les paramètres gestuels évoqués ? Cette problématique sera abordée dans le chapitre 5 dans le cadre du projet ANR MetaSon : Métaphores Sonores. Des études réalisées en collaboration avec le Laboratoire de Neurosciences Cognitives de Marseille (LNC - UMR 7291) seront présentées. Ces études ont pour objectifs de mettre en évidence la faisabilité de tels outils de guidage des paramètres dynamiques d'un geste d'écriture afin d'envisager des applications pour la réhabilitation d'une pathologie graphomotrice : la dysgraphie.

Chapitre 2

Du son aux formes géométriques : perception auditive des mouvements graphiques

Sommaire

2.1	Introduction	49
2.2	Synthèse sons de frottement	51
2.3	Expérience 1 - Reconnaissance auditive des mouvements biologiques	51
2.4	Expérience 2 - Discrimination auditive de formes géométriques distinctes	54
2.5	Expérience 3 - Discrimination auditive de formes géométriques proches	61
2.6	Le profil de vitesse : un invariant perceptif	64
2.7	Discussion générale	67

L'étude présentée dans ce chapitre est issue de l'article de journal :

Thoret, E., Aramaki, M., Kronland-Martinet, R., Velay, J. L., & Ystad, S. (2014). From Sound to Shape : Auditory Perception of Drawing Movements, *Journal of Experimental Psychology : Human Perception and Performance*, 40, 983-994. doi :10.1037/a0035441

et du chapitre d'ouvrage :

Thoret, E., Aramaki, M., Kronland-Martinet, R., Velay, J. L., & Ystad, S. (2013). Reenacting sensorimotor features of drawing movements from friction sounds. In *From Sounds to Music and Emotions*, LNCS vol 7900, 130-153. Springer Berlin Heidelberg. doi :10.1007/978-3-642-41248-6_8

2.1 Introduction

Les travaux présentés dans cette première étude se sont donc focalisés sur la perception auditive d'un type particulier de mouvement, les mouvements graphiques, spécifiques aux humains. Dans la mesure où ces mouvements génèrent une trace visuelle, mais aucun son spécifique, ces mouvements sont plus naturellement associés à la vision qu'à l'audition. Cependant, si on prête attention aux sons de frottement produits quand une personne dessine, on peut entendre très distinctement le son du frottement sur le papier, en particulier quand la surface est très rugueuse. Comme nous ne prêtons

pas spécialement attention à ces sons, il semble *a priori* improbable que nous soyons capables de reconnaître une forme dessinée juste sur la base des sons de frottement produits quand une personne dessine. Cependant, ces sons représentent plus qu'un simple bruit. En effet, les changements de vitesse du stylo produisent des variations de timbre qui peuvent peut-être permettre de s'imaginer le geste réalisé par le scripteur. Dans ce premier chapitre, nous avons donc souhaité évaluer dans quelle mesure il est possible d'inférer ce qui est dessiné juste à partir du son. En particulier, nous avons tenté d'identifier les paramètres acoustiques qui permettent de reconnaître un mouvement et de plus s'ils permettent de remonter aux caractéristiques géométriques d'une forme.

Ces problèmes ont été étudiés à travers trois expériences. Dans un premier temps, nous avons voulu vérifier qu'il était possible de reconnaître les variations de vitesse réalisées quand une personne dessine juste à partir des sons de friction produits par le stylo. A cette fin, nous nous sommes basés sur des travaux séminaux de Viviani et ses collègues sur la perception visuelle des mouvements biologiques et ses relations avec le système moteur. Ces travaux ont mis en évidence que quand nous produisons nos mouvements, la cinématique de nos gestes était reliée à la courbure de la trajectoire parcourue (Viviani et Terzuolo, 1982) et qu'en particulier la vitesse tangentielle est proportionnelle à la puissance $1/3$ de la courbure de la trajectoire (Lacquaniti *et al.*, 1983; Viviani et Flash, 1995; Viviani et McCollum, 1983).

Dans une première expérience (Expérience 1), nous avons adapté un protocole utilisé dans le domaine visuel par Viviani et Stucchi (1992) à la perception auditive des sons de frottement. Les sujets devaient manipuler la loi en puissance qui contrôlait un son de frottement afin que le son leur évoque le geste le plus réaliste et naturel possible. Cette expérience a été rendue possible par l'utilisation de sons de frottement synthétiques dont la génération a été présentée dans le chapitre 1. Notre première hypothèse était donc que les variations de timbre générées par les variations de vitesse devraient permettre de reconnaître la valeur spécifique de l'exposant de la loi en puissance reliant vitesse et courbure. Ainsi, les variations de timbre d'un geste humain pourraient être considérées comme un invariant transformationnel permettant la reconnaissance des gestes humains par le biais du son. Dans un second temps, nous avons émis l'hypothèse que si ce type de sons permettait de reconnaître un geste, ils devaient permettre de reconnaître, dans une certaine mesure la forme géométrique qui était dessinée. Dans deux expériences (Expériences 2 et 3), nous avons donc demandé à des sujets d'associer les sons de frottement qui étaient produits aux formes statiques qui étaient dessinées. Afin d'évaluer quelle était l'information gestuelle utilisée pour réaliser une éventuelle reconnaissance de ces formes, nous avons utilisé à la fois des sons de frottement enregistrés et des sons synthétisés uniquement à partir du profil de vitesse de la personne dessinant les formes. Dans l'expérience 2, les formes étaient aisément discriminables, à la fois du point de vue perceptif et cinématique. Dans l'expérience 3, les formes à discriminer étaient supposées être perceptivement plus proches car leurs profils de vitesse étaient plus similaires que ceux des formes de l'expérience 2. Notre hypothèse était donc que des formes plus proches du point de vue cinématique devraient impliquer plus de confusions.

Les résultats de ces trois expériences, et en particulier la pertinence du profil de vitesse du point de vue perceptif, seront discutés dans le cadre de l'approche écologique et de la taxonomie des invariants présenté dans le chapitre 1. Ces résultats seront également partiellement discutés dans le contexte de l'approche idéomotrice utilisée par Young *et al.* (2013), afin de discuter la perception des sons en termes de représentations cognitives résultant d'un couplage structurel entre le flux sensoriel et son traitement. Avant de présenter en détail ces trois expériences, la section suivante décrit

succinctement les principes génériques de la synthèse sonore de sons de frottement.

2.2 Synthèse sons de frottement

La synthèse sonore est un outil particulièrement approprié pour étudier la perception du geste évoqué par un son de frottement et nous avons présenté dans le chapitre 1 un modèle permettant de générer des sons de frottement à partir d'un profil de vitesse spécifique, ce modèle va être rappelé rapidement ici. En effet, la possibilité de manipuler la dynamique du timbre à partir d'un profil de vitesse donné permet d'envisager la construction de protocoles expérimentaux originaux. Gaver (1993a) et Van Den Doel *et al.* (2001) ont ainsi proposé un modèle de synthèse physiquement informé de sons de frottement permettant de générer un son de frottement de synthèse uniquement à partir d'un profil de vitesse. Ce modèle simule le son d'une source sonore résultant d'une suite d'impacts d'un plectre sur les aspérités d'une surface. La rugosité de la surface est modélisée par un bruit blanc représentant les hauteurs des aspérités de la surface. Le son de synthèse peut ainsi être généré en filtrant passe-bas le bruit et en contrôlant la fréquence de coupure du filtre par la vitesse du plectre, et donc ici par la vitesse du geste.

En plus du profil de vitesse, d'autres paramètres du modèle de synthèse tels que la pression ou encore l'angle du stylo pourraient être considérés dans un modèle plus sophistiqué. Par exemple, la pression du stylo pourrait être reliée à l'intensité du son. Cependant, le but de cette étude est d'évaluer l'influence du profil de vitesse sur la perception du geste sous jacent à un son. Dans la suite de l'étude, la pression sera donc supposée constante durant toute la durée du mouvement. Ainsi, seul le profil de vitesse du geste sera transmis par les variations de timbre du son de frottement.

2.3 Expérience 1 - Reconnaissance auditive des mouvements biologiques

Afin d'évaluer notre capacité à reconnaître des gestes spécifiques à travers le son, nous nous sommes basés sur un protocole de Viviani et Stucchi (1992) concernant la perception visuelle des mouvements biologiques. Dans cette expérience, les sujets devaient ajuster la vitesse d'un point lumineux afin de percevoir le mouvement le plus uniforme, i.e. la vitesse la plus constante possible. Le mouvement du point respectait la loi en puissance suivante :

$$v_t(s) = KC^{-\beta} \quad (2.1)$$

où v_t est la vitesse tangentielle du mouvement, C la courbure et K le gain de vitesse relié à la vitesse moyenne du mouvement (Viviani et Terzuolo, 1982; Lacquaniti *et al.*, 1983; Viviani et McCollum, 1983; Viviani et Flash, 1995). La tâche était implicite et les sujets n'étaient pas conscients qu'ils ajustaient l'exposant β de la loi en puissance. Les résultats ont alors mis en évidence que la vitesse perçue comme la plus uniforme correspondait à un ajustement de l'exposant de la loi à une valeur proche d'un tiers en moyenne, valeur correspondant à celle obtenue pour un mouvement graphique réel. A l'inverse, quand l'exposant était nul, ce qui correspond une vitesse constante quelle que soit la courbure de la trajectoire, le mouvement du point était perçu comme accélérant dans les parties les plus courbées de la trajectoire. Des simulations numériques ont également mis en évidence que l'exposant $1/3$ minimise le *jerk*, grandeur caractérisant l'aspect saccadé du mouvement (Richardson et Flash, 2002; Viviani et Flash, 1995).

Dans cette étude, une expérience similaire à celle de Viviani et Stucchi (1992) a été mise en place du point de vue sonore. Le point lumineux utilisé pour étudier la perception visuelle du mouvement était ici remplacé par un son de frottement de synthèse évoquant une vitesse par ses variations intrinsèques de timbre. Les sujets devaient ajuster implicitement l'exposant de la loi en puissance. Il


est important de noter que du point de vue visuel, Viviani et ses collègues étudiaient la relation entre la courbure d'une forme géométrique et la cinématique du point lumineux, tandis que les sons monophoniques utilisés ici n'évoquent que la cinématique du geste déterminée par le profil de vitesse tangentielle. Si nos hypothèses sont correctes, les sons de frottement doivent évoquer, par le biais de leurs variations de timbre, la vitesse du mouvement de la source physique, c'est-à-dire l'accélération du stylo dans les parties les plus plates de la trajectoire et inversement sa décélération dans les parties les plus courbées.

2.3.1 Méthode

Participants

20 participants volontaires âgés en moyenne de 29.42 ans (SD = 12.54) ont pris part à cette expérience (3 femmes). Aucun des sujets n'était au courant des objectifs de la recherche avant de participer à l'expérience.

Stimuli

 Les sons de frottement étaient synthétisés en utilisant le modèle de synthèse sonore décrit précédemment. Afin d'éviter d'évoquer des formes géométriques spécifiques, et ainsi ne considérer que le geste évoqué, les sons de frottement étaient générés à partir de trajectoires pseudo-aléatoires d'un point mobile $(x(t), y(t))$ défini par le système d'équations paramétriques suivant :

$$\begin{cases} x(t) = A_0 \sum_{k=0}^3 a_k \sin(\omega_{x,k} t) \\ y(t) = B_0 \sum_{k=0}^3 b_k \sin(\omega_{y,k} t) \end{cases} \quad (2.2)$$

avec $A_0 = 7$, $B_0 = 5$, $a_0 = 1$, $b_0 = 1$.

Le reste des paramètres était régénéré aléatoirement toutes les 15 secondes. En particulier, les valeurs de $\omega_{x,k}$ et $\omega_{y,k}$ étaient tirées aléatoirement entre 0 et 0.6 Hz et a_k et b_k entre 0 et 0.9 Hz. De cette façon, le mouvement du point n'était *prévisible* que pendant 15 secondes. Le profil de vitesse utilisé pour générer les sons était calculé en utilisant la loi en puissance (Eq. 2.1) avec $K = 10 \text{ m.s}^{-1}$. La courbure C était calculée à partir de la formule suivante :

$$C(t) = \frac{|\ddot{x}\dot{y} - \ddot{y}\dot{x}|}{(\dot{x}^2 + \dot{y}^2)^{3/2}} \quad (2.3)$$

où x et y sont définis selon l'équation 2.2, et \dot{x} , \ddot{x} (respectivement \dot{y} , \ddot{y}) représentent la dérivée et la dérivée seconde par rapport au temps de la variable x (respectivement y). Le son de friction associé était synthétisé en temps-réel à partir de la valeur de l'exposant de la loi en puissance (Eq. 2.1). La valeur de l'exposant était modifiée directement par les sujets durant l'expérience.

Procédure

Les participants étaient assis devant un écran d'ordinateur dans une cabine audiométrique. Les sons étaient présentés par le biais d'un casque d'écoute Sennheiser HD-650. L'expérience était précédée d'une session d'entraînement avec 2 essais. A chaque essai, une forme pseudo-aléatoire différente était générée et le son de synthèse associé était généré en temps réel. Les trajectoires étaient différentes à travers les essais et les sujets. Il est important de noter que la trajectoire n'était pas présentée aux sujets et que le test était uniquement sonore.

Il était demandé aux sujets d'imaginer que le son qu'ils entendaient était produit par quelqu'un en train de dessiner une forme aléatoire, un *gribouilli* par exemple, sur une surface rugueuse. Il leur était demandé d'ajuster le son en utilisant deux boutons, "<" et ">", affichés sur l'écran de l'ordinateur, jusqu'à ce que le son qu'ils percevaient leur évoque le geste le plus *naturel*, le plus *fluide*. L'interface graphique a été créée à partir du logiciel Max/MSP¹. Les sujets n'étaient pas au courant de la façon dont les boutons, qui en réalité ajustaient l'exposant de la loi en puissance, modifiaient le son. Les sujets disposaient du temps qu'ils souhaitaient pour réaliser les ajustements. Il leur était vivement conseillé de parcourir l'ensemble des sons possibles. Les valeurs de l'exposant variaient entre 0 et 1.0816 avec un pas constant de 0.0416. La valeur initiale de l'exposant était choisie aléatoirement au début de chaque essai. Pour chaque sujet, la valeur finale de l'exposant de chacun des 6 essais était enregistrée.

2.3.2 Résultats

Les valeurs finales de l'exposant étaient moyennées à travers les 6 essais, dans un premier temps par sujet puis à travers les sujets. La valeur moyenne de l'exposant à travers les sujets est de 0.361 (SD=.084). Cette valeur ne diffère pas significativement d'un tiers ($t(19) = 1.53; p = .14; d = .329$ – t -test bilatéral à un échantillon). Une analyse des résultats en fonction des valeurs initiales a été réalisée et une corrélation significative entre valeurs initiales et valeurs finales a été trouvée ($r = .34, p < .05$). Comme proposé dans une étude de Carlyon *et al.* (2010), nous avons effectué une analyse complémentaire afin de nous assurer que les valeurs de l'exposant ont bien été ajustées autour d'une valeur proche d'un tiers. Deux groupes d'exposants finaux ont été considérés : les essais dont la valeur initiale était inférieure à 1/3 et ceux dont la valeur finale était supérieure à 1/3. Au final, à travers tous les sujets, 70 valeurs initiales étaient supérieures à 1/3 et 50 étaient inférieures. Une valeur finale de l'exposant de .31 (95% CI [0.28;0.34]) a été trouvée pour les seuils ascendants, et une valeur finale de .39 (95 % [0.36;0.43]) pour les seuils descendants. De plus, afin de s'assurer qu'un ajustement a bien été effectué, les écarts-types des valeurs initiales et finales ont été comparées. Un t -test unilatéral à deux échantillons a révélé que les écarts-types finaux étaient significativement plus petits que les écarts-types des valeurs initiales, $t(19) = -7.86; p < .05$.

2.3.3 Discussion

Dans cette expérience, les sujets devaient modifier le timbre du son qu'ils entendaient afin que ses variations leur évoquent le geste humain le plus naturel et fluide possible. Il leur était précisé que le son était produit par quelqu'un qui dessinait une forme sans géométrie spécifique sur une feuille en papier. Le son était généré par synthèse en utilisant le profil de vitesse calculé à partir de l'équation 4.5. Les sujets ajustaient implicitement la valeur de l'exposant de la loi en puissance et n'étaient pas au courant de la façon dont leurs réglages influent sur le son.

Le résultat principal de cette expérience est que pour évoquer par le son le geste d'une personne qui dessine une forme pseudo-aléatoire de la façon la plus naturelle et fluide possible, les sujets ajustent l'exposant autour d'une valeur proche d'un tiers. Dit autrement, cette expérience montre que pour que le timbre du son de frottement d'une personne qui dessine évoque un geste naturel, il doit varier en adéquation avec les règles de vrais mouvements graphiques, i.e. un mouvement biologique (Viviani et Terzuolo, 1982). Ce résultat révèle que nous sommes capables de nous imaginer un geste humain naturel, i.e. les mouvements biologiques, à partir de la modalité auditive et en particulier des variations de timbre d'un son. Ce résultat est à mettre en perspective avec la littérature sur la reconnaissance des événements sonores grâce à des pattern acoustiques spécifiques, aussi appelés

1. <http://cycling74.com>

invariants transformationnels et caractérisant l'action effectuée dans le son. Dans une perspective écologique, ce résultat met en évidence que le profil de vitesse des gestes, et en particulier la loi en puissance $1/3$, est caractéristique des mouvements biologiques au niveau sonore et peut donc être considéré comme un invariant transformationnel de la perception auditive.

Dans cette première expérience les sujets étaient informés que la forme qui était dessinée correspondait à des gribouillis et il leur était demandé uniquement d'inférer une dynamique gestuelle à partir du son. Dans la mesure où la loi en puissance $1/3$ relie la vitesse d'un geste à la courbure de la forme dessinée, et donc implicitement à sa géométrie, il est envisageable d'évoquer la géométrie d'une trajectoire particulière à partir des sons de frottement comme cette loi est reconnaissable du point de vue sonore. En effet, un tel processus d'association entre un son et une forme pourrait s'appuyer sur le fait que si on incarne cette loi, on incarne le principe suivant : *les hautes vitesses évoquées correspondent à des petites courbures et les faibles vitesses à des grandes courbures*. Bien que le son de synthèse utilisé dans cette expérience ne transmette qu'une seule variable, le profil de vitesse, la connaissance implicite de la loi en puissance $1/3$ devrait permettre d'inférer une description géométrique de la forme dessinée juste à partir du son.

2.4 Expérience 2 - Discrimination auditive de formes géométriques distinctes

Une forme visuelle peut être décrite à la fois par sa géométrie et, en terme moteur, par le mouvement requis pour la dessiner. Deux études séminales de Freyd (1983a,b) ont mis en évidence que la perception d'une forme visuelle était intimement liée à sa représentation motrice. En particulier, elle a mis en évidence qu'une personne qui lit une lettre utilise ses connaissances motrices afin d'en décoder son contenu visuel *statique*. Elle a même montré qu'une relation fonctionnelle très forte existait entre les caractéristiques visuelles *statiques* d'une forme géométrique et le mouvement graphique nécessaire pour la réaliser. Par ailleurs, différentes études ont mis en évidence que voir une forme graphique activait les mêmes processus cérébraux que ceux requis pour la dessiner si les sujets savaient déjà la dessiner (James et Gauthier, 2006; Longcamp *et al.*, 2003, 2006). De plus, Longcamp *et al.* (2008) ont mis en évidence que les mouvements spécifiques qui doivent être réalisés pour dessiner un caractère nouveau sont mémorisés afin d'être utilisés dans le processus de reconnaissance visuel de ce même caractère. Ces résultats sont à mettre en regard avec une étude de Viviani et Stucchi (1989) qui a montré qu'une forme décrite par un point lumineux en mouvement pouvait être aisément reconnue grâce à la cinématique du point lumineux si celle-ci respectait celle des mouvements biologiques. À l'inverse, cette même étude a mis en évidence que la perception dynamique de la géométrie d'une telle forme était distordue quand la cinématique du point lumineux ne respectait pas la loi en puissance $1/3$, i.e. les mouvements biologiques. Par exemple, il a montré que si un cercle était parcouru par un point lumineux qui accélérât dans ses parties opposées, il était perçu plus aplati que s'il était parcouru à vitesse constante, respectant ainsi la loi en puissance $1/3$.

À l'inverse de la vision, aucun son n'est naturellement associé à une forme géométrique car une forme géométrique est un objet abstrait associé à aucune source sonore physique. L'expérience précédente a mis en évidence notre capacité à reconnaître la dynamique des mouvements biologiques, i.e. la loi en puissance $1/3$, par le biais des sons de frottement produits lorsqu'une personne dessine. Dans la mesure où cette relation relie les caractéristiques dynamiques du geste, i.e. sa vitesse, aux caractéristiques géométriques, i.e. sa courbure, on peut faire l'hypothèse que la trajectoire parcourue par le geste est elle aussi intégrée implicitement en écoutant le son de frottement.

Dans cette seconde expérience, notre but est d'évaluer dans quelle mesure nous sommes capables de reconnaître une forme géométrique par le biais des sons de frottement. Nous avons ainsi demandé à des sujets d'associer les sons de frottement produits à la forme statique qu'ils pensaient être celle dessinée. Les sujets devaient discriminer quatre formes géométriques sur la base des quatre sons de frottement générés lors de leur production. Le but d'une telle tâche de discrimination était de limiter le nombre d'associations possibles entre les sons et les formes. Au-delà d'évaluer notre capacité à reconnaître une forme par le biais du son, notre second but était aussi de déterminer quelle était l'information contenue dans les sons qui permettait d'effectuer une telle tâche. Nous avons donc utilisé deux types de sons, à savoir les sons produits par une personne qui dessinait les quatre formes simples et les sons re-synthétisés à partir des profils de vitesse enregistrés sur une tablette graphique. En effet, en plus de la cinématique, un son de friction enregistré contient des indices acoustiques subtiles qui peuvent varier selon d'autres paramètres que la vitesse du geste. L'angle du stylo, la pression à laquelle le stylo appuie sur la feuille, et les irrégularités de la surface peuvent en effet fournir d'autres indices acoustiques permettant la discrimination de formes géométriques par le biais du son. De plus, ces indices peuvent être reliés aux pleins et aux déliés effectués quand une personne produit une lettre et donc directement reliés à la représentation gestuelle. Afin de tester l'hypothèse que le profil de vitesse seul est suffisant pour reconnaître une forme géométrique par le biais du son, les résultats issus des deux corpus sonores ont été analysés et comparés. En s'appuyant sur les résultats de l'expérience précédente, on souhaitait donc également vérifier l'hypothèse que la cinématique du geste, par le biais de la loi en puissance $1/3$, favorisait une telle association entre son et forme.

2.4.1 Méthode

Participants

20 participants volontaires d'âge moyen de 30.65 ans ($SD = 13.11$) ont pris part à cette expérience (9 femmes). 12 sujets étaient naïfs aux objectifs de la recherche avant de participer à l'expérience. 8 sujets avaient également participé à l'Expérience 1.

Stimuli

Formes statiques. Des tests préliminaires informels ont mis en évidence que les formes géométriques qui possèdent des points de rebroussement étaient facilement reconnaissables, en effet, ces particularités géométriques causent une discontinuité dans le son, i.e. *un silence*. Basé sur ces observations préliminaires, un corpus de quatre formes a été réalisé, deux avec des points de rebroussement (des arches, un trait), et deux sans (un cercle, une ellipse). Les formes sans point de rebroussement, i.e. cercle et ellipse, diffèrent uniquement par leur excentricité, tandis que les arches et les traits, diffèrent par la position des points de rebroussement sur la forme. La figure 2.2(a) présente les quatre formes.

Sons enregistrés. 🎧 Les sessions d'enregistrements ont eu lieu dans un studio. Un membre de l'équipe a dessiné les quatre formes de la façon la plus fluide possible sur une tablet Wacom Intuos3. A part pour les arches, il lui était demandé de ne pas lever le stylo de la tablette pendant 25 secondes pour être certains que plusieurs périodes étaient produites convenablement. Pour les arches, il lui était demandé de parcourir la tablette de gauche à droite jusqu'à ce qu'il atteigne la fin de la forme et de retourner à la position initiale de la forme sans lever le stylo. Les mouvements du scripteur étaient enregistrés sur la tablette avec une fréquence d'échantillonnage temporel de 200 Hz et une précision spatiale de 5.10^{-3} mm. La figure 2.1 présente le dispositif d'enregistrement. En plus

TABLE 2.1 – Caractéristiques géométriques et temporelles des séquences choisies pour l’expérience 2

Forme	Longueur théorique (cm)	Longueur (cm)	Durée (secondes)
Cercle	62.02	62.5	5.2
Ellipse	92.04	89.32	5.8
Arches	82.01	87.11	5.1
Trait	80	88.7	5.2

de l’enregistrement du mouvement, des enregistrements monophoniques des sons produits ont été réalisés avec une fréquence d’échantillonnage de 44100 Hz à l’aide d’un micro cardioïde Neumann-KM84i positionné à environ 30 cm de la tablette. Des séquences de chaque enregistrement ont été sélectionnées pour l’expérience en fonction de la régularité géométrique et temporelle des productions du scripteur. En particulier, des séquences de durées similaires (environ 5 secondes) d’un même nombre de périodes ont été choisies. Pour les ellipses, arches, et traits, les séquences sélectionnées correspondaient à 4 périodes. Pour les cercles, seulement deux périodes ont été considérées car la durée moyenne d’une période d’un cercle était environ le double de celle des trois autres formes. La table 2.1 résume les caractéristiques temporelles et géométriques des quatre formes sélectionnées pour l’expérience. Un fenêtrage d’environ 10 ms a été utilisé pour la segmentation afin d’éviter les *clics* au début et à la fin de chaque séquence sonore. Les sons ont été normalisés à -3 dB. Les profils de vitesse correspondant aux séquences ont été calculés à partir des enregistrements de la tablette graphique pour les quatre formes. La figure 2.2(b) présente une période du profil de vitesse de chaque forme après normalisation en amplitude et filtrage passe-bas à 10 Hz. Les passages par zéro de la vitesse sont signalés par des ronds noirs. Ils est intéressant de noter que les quatre formes ont des profils de vitesse très différents, ce qui nous permet de faire l’hypothèse qu’ils sont plus facilement discriminables du point de vue perceptif.



FIGURE 2.1 – Dispositif utilisé pour enregistrer les sons et les profils de vitesse produit par le scripteur

Sons de synthèse. © Les sons de frottement synthétiques ont été générés en utilisant le même modèle que dans l’expérience 1. Pour les quatre formes, les profils de vitesse utilisés sont ceux qui ont été enregistrés à l’aide de la tablette graphique durant les séances d’enregistrement.

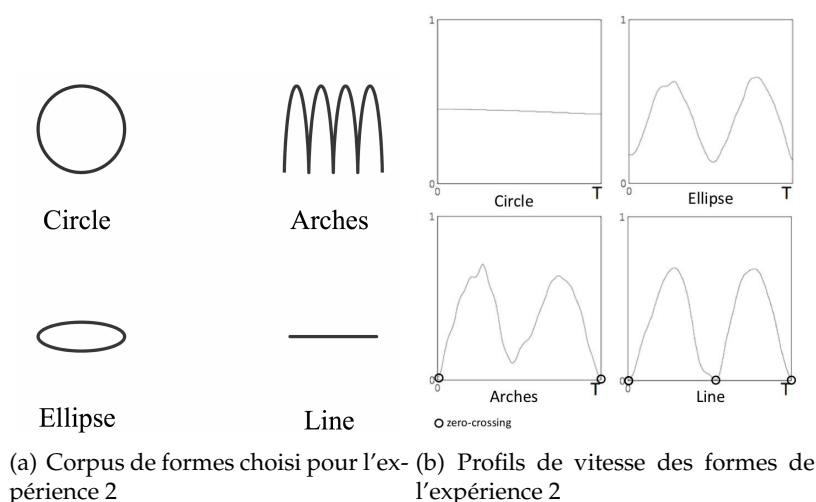


FIGURE 2.2 – Formes et profils de vitesse associés de l'expérience 2 – figure issue de Thoret *et al.* (2014c).

Tâche

Les participants étaient assis devant un écran d'ordinateur dans une cabine audiométrique. Les sons étaient présentés par le biais d'un casque d'écoute Sennheiser HD-650. L'interface graphique du test a été programmée avec le logiciel Max/MSP. L'expérience comprenait huit essais, quatre pour les sons de synthèse et quatre pour les sons enregistrés. Les essais étaient présentés dans un ordre aléatoire à travers les participants. A chaque essai, quatre icônes représentant les sons étaient affichées à gauche de l'écran de l'ordinateur. Les sujets avaient pour consigne d'associer chaque son à la forme qui correspondait au mieux, c'est-à-dire la forme évoquée par le son produit par le scripteur. Les formes étaient toujours disposées de la même façon alors que les sons, tous représentés par le même icône, étaient présentés dans une position aléatoire à travers les essais (voir figure 2.3). Les sujets pouvaient glisser-déposer les sons à côté des formes. Ils savaient que tous les sons avaient été enregistrés pendant qu'un scripteur dessinait les formes de façon fluide, sans lever le stylo. Aucune limitation de temps n'a été imposée et les sujets pouvaient écouter les sons autant de fois qu'ils le souhaitaient. A chaque essai, quatre associations son/forme étaient collectées.

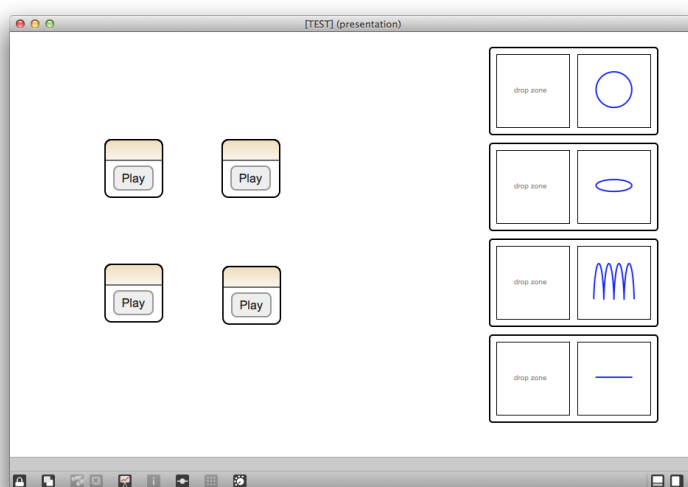


FIGURE 2.3 – Interface du test.

Analyse des données

Les associations entre un son et une forme étaient jugées à 1 quand le son était associé la bonne forme et à 0 sinon. Ainsi, pour chaque sujet et pour chaque type de son (synthétisé et enregistré), les valeurs étaient moyennées à travers les essais afin d'établir une matrice d'association contenant les scores de la tâche d'association. Dans la suite, on appellera taux de réussite le pourcentage d'association son/forme correcte, ce qui correspond aux valeurs de la diagonale de la matrice d'association.

Afin d'analyser les erreurs des sujets, les taux d'associations entre un son donné et les quatre formes, puis entre une forme et les quatre sons, ont été comparés. A cette fin, des *t*-tests pour deux échantillons appariés ont été réalisés entre les taux de réussite et les six autres taux d'association. En plus de ces comparaisons par paire, un test de conformité à un standard a été réalisé avec un seuil correspondant à un taux d'association de 25%, taux correspondant à la chance statistique.

Afin d'évaluer si la tâche a été réalisée de la même façon pour les sons enregistrés que pour les sons synthétisés, et ainsi dans le but de voir si le profil de vitesse est une information perceptive suffisante pour la discrimination de forme, une comparaison globale entre les matrices d'association pour les deux types de sons a également été réalisée. Pour cela, un test de corrélation de Pearson sur les distances cophénétiques entre les formes de chaque matrice a été réalisé (une analyse détaillée des comparaisons entre les matrices d'association est présentée en section 2.6). En pratique, le test a été réalisé de la façon suivante : pour chaque type de son, une matrice de *dissociation* D^* a été calculée dans un premier temps, définie par $D^* = 1 - S^*$, où S^* est la version symétrisée de la matrice d'association S . Une matrice D^* de distance deux à deux a ainsi été calculée depuis D en utilisant une métrique euclidienne. Une analyse en cluster hiérarchique de D (complete linkage) a ensuite été réalisée. Les distances cophénétiques ont finalement été calculées depuis le dendrogramme résultant de cette analyse et ont été assemblées en un vecteur. Il est ici nécessaire d'introduire la notion de distance cophénétique.

Le problème de la comparaison entre deux dendrogrammes est ancien et a été abordé dans le domaine de la phylogénétique. Un des objectifs de ce domaine de la biologie est de comprendre l'évolution des espèces selon des critères à la fois génétiques, et donc objectifs, et selon des critères morphologiques, qui relèvent plus d'une observation empirique. Ces deux types d'observations permettent d'établir des distances entre les différentes espèces et ainsi d'observer ces données sous la forme de dendrogrammes. Un dendrogramme est une représentation des distances entre différents objets, et la composition de ses différents clusters est faite selon une métrique spécifique. Un dendrogramme peut donc ainsi être caractérisé par les distances entre ses différents clusters, appelée la distance cophénétique (Sokal et Rohlf, 1962). Afin de clarifier cette notion, la figure 2.4 présente un dendrogramme et les distances cophénétiques associées. Les distances cophénétiques sont stockées dans un tableau pour chaque dendrogramme. On peut donc comparer deux dendrogrammes en calculant les corrélations de Spearman et de Pearson. La corrélation de Pearson permet d'effectuer une comparaison quantitative de la corrélation linéaire entre les deux formes. La corrélation de Spearman permet elle d'effectuer une comparaison qualitative des deux vecteurs de distances qui tient compte du rang des distances cophénétiques, i.e. leur position relative sur l'axe des abscisses. Ainsi, pour comparer les deux matrices obtenues pour les sons de synthèse et les sons enregistrés, les deux vecteurs ont été soumis à un test de corrélation de Pearson.

Enfin, afin de comparer plus précisément les taux de réussite entre les sons enregistrés et synthétisés, un test bilatéral de Wilcoxon à rangs signés a été réalisé sur les taux de réussite pour chaque forme avec le type de son comme facteur.

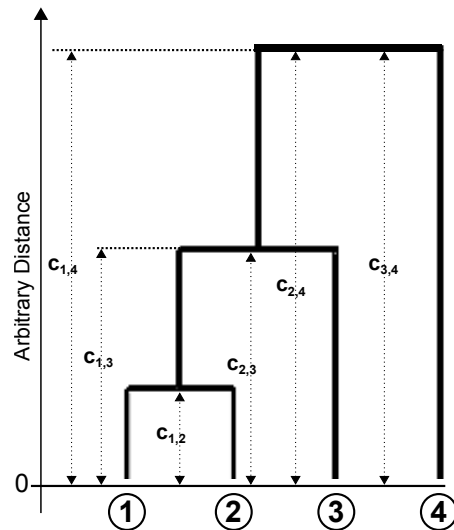


FIGURE 2.4 – Exemple de dendrogramme et des distances cophénétiques associées $c_{i,j}$ entre différents objets – Les distances cophénétiques sont résumées dans le tableau suivant : $C = [c_{1,2}, c_{1,3}, c_{1,4}, c_{2,3}, c_{2,4}, c_{3,4}]$ – et peuvent être comparées avec les distances cophénétiques d’autres dendrogrammes avec des corrélations de Pearson et Spearman – Figure issue de Thoret *et al.* (b).

Dans toutes les analyses statistiques qui suivent le risque a été défini à .05. Toutes les comparaisons ont été ajustées avec les corrections de Bonferroni pour les comparaisons multiples.

2.4.2 Résultats

La table 2.2 présente les matrices d’associations moyennées à travers les sujets pour chaque type de son. Pour tous les sons, les taux de réussite (i.e. la diagonale de la matrice) sont significativement supérieurs à la chance ($p < .001$ pour chaque forme, $d_{cercle} = 13.19$, $d_{ellipse} = 2.01$, $d_{arches} = 1.91$, $d_{trait} = 9.42$ pour les sons enregistrés, $d_{cercle} = 13.19$, $d_{ellipse} = 2.81$, $d_{arches} = 2.23$, $d_{trait} = 4.9$ pour les sons enregistrés) et en particulier tous supérieurs à 80% avec des taux de réussite les plus élevés pour le trait et le cercle (proche de 100%). De plus les taux de réussite diffèrent significativement des trois autres taux d’association ($p < .001$ pour toutes les comparaisons). Ces premiers résultats révèlent que les sons de frottement ont été associés correctement aux formes géométriques associées.

L’analyse des distances cophénétiques met en évidence que les matrices d’association pour les sons enregistrés et synthétisés sont fortement corrélés, $r(4) = .89$, $p < .05$. De plus, les tests de Wilcoxon mettent en évidence que les taux de réussite ne diffèrent pas significativement entre les sons de synthèse et les sons enregistrés (Cercle : $z = 0$, $p = 1$; Ellipse : $z = -1.265$, $p = .21$; Arches : $z = -.632$, $p = .52$; Trait : $z = -1.13$, $p = .25$). Cela révèle que les sons de frottement de synthèse et les sons enregistrés ont des taux d’associations comparables.

2.4.3 Discussion

Les résultats de cette expérience montrent que les sujets ont été capables d’associer un son de frottement (choisis parmi quatre) à la forme qui correspondait. Les taux de réussite sont élevés pour chaque forme. De plus les taux d’association obtenus pour les sons de synthèse et les sons enregistrés ne diffèrent pas significativement bien que les sons de synthèse ne contiennent que l’information cinématique du mouvement du scripteur. Les deux types de sons ne diffèrent en fait que par le fait que d’autres variables, telles que la pression, les irrégularités de la surface, ou encore l’angle du stylo, ont pu apporter d’autres indices acoustiques dans les sons enregistrés. Cependant, en dépit de ces indices acoustiques supplémentaires, les taux de réussite pour les sons enregistrés ne se révèlent

TABLE 2.2 – Matrices d’association sons/formes – Moyennes et erreurs standards en pourcentage – Expérience 2. Les taux d’association supérieurs à la chance (25%) sont notés en gras. Le niveau de significativité des comparaisons à la chance est noté : * $p < .05$, ** $p < .01$, *** $p < .001$

Forme	Sons enregistrés			
	Cercle	Ellipse	Arches	Trait
Cercle	98.75***	1.25	-	-
SE	2.5	2.5	-	-
Ellipse	1.25	81.25***	17.5	-
SE	2.5	8.78	8.55	-
Arches	-	17.5	80***	2.5
SE	-	8.55	9	3.51
Trait	-	-	2.5	97.5***
SE	-	-	3.51	3.51

Forme	Sons synthétisés			
	Cercle	Ellipse	Arches	Trait
Cercle	98.75***	1.25	-	-
SE	2.5	2.5	-	-
Ellipse	-	87.5***	12.5	-
SE	-	7.44	7.44	-
Arches	-	11.25	82.50***	6.25
SE	-	7.11	8.55	5.45
Trait	1.25	-	5	93.75***
SE	2.5	-	4.9	5.45

pas être plus grands que pour les sons de synthèse. Ce résultat est à mettre en regard avec une étude de Schomaker et Plamondon (1990) qui a montré qu'aucune relation générale, telle que la loi en puissance $1/3$ pour la vitesse, n'existe entre ces paramètres gestuels (pression et angle) et la géométrie de chaque forme. De plus, comme les sujets qui ont passé le test n'avaient pas dessiné les formes eux-mêmes, ils n'ont pas pu établir de relation entre des potentiels indices acoustiques liés à ces paramètres gestuels et la géométrie, ce qui aurait pu éventuellement permettre d'augmenter encore plus le taux de reconnaissance pour les sons enregistrés.

Ce résultat soutient le fait que le profil de vitesse constitue l'information principale utilisée pour la reconnaissance de formes géométriques par le biais du son. En particulier, les événements singuliers tels que les silences, correspondant à des vitesses nulles, ont pu être la cause des forts taux de reconnaissance. Du point de vue cognitiviste, on peut donc faire l'hypothèse que notre capacité à associer des sons à une forme géométrique se base sur un modèle interne du geste évoqué par la perception visuelle de la forme et la perception auditive des variations de timbre évoquant le profil de vitesse du geste.

Enfin, dans cette expérience, les forts taux de réussites ont été obtenus pour un corpus de formes relativement distinctes compte tenu de la présence ou non de point de rebroussement et par conséquent du geste produit pour dessiner ces formes. Dans l'expérience suivante, notre but était d'évaluer si les associations entre sons de frottement et formes géométriques étaient encore possible pour des formes plus proches du point de vue géométrique, et aussi du point de vue gestuel, c'est-à-dire sans la présence de points de rebroussement créant des indices acoustiques (i.e. des silences) très saillants.

2.5 Expérience 3 - Discrimination auditive de formes géométriques proches

2.5.1 Méthode


La procédure et l'analyse des données étaient les mêmes que pour l'expérience 2.

Participants

18 participants volontaires d'âge moyen de 31.56 ans ($SD = 13.73$) ont pris part à cette expérience (8 femmes). 11 sujets étaient naïfs aux objectifs de la recherche avant de participer à l'expérience. 7 sujets ont également participé à l'expérience 1 et 17 à l'expérience 2.

Stimuli

Formes statiques. Comme pour l'expérience précédente, des tests informels ont été réalisés afin de choisir le corpus de formes géométriques et les sons à partir de critères perceptifs. En particulier, les formes choisies ne présentaient plus de point de rebroussement, dans l'hypothèse où elles seraient moins facilement discriminables du point de vue perceptif. En conséquence, le cercle et l'ellipse ont été conservés pour cette seconde expérience, tandis que le trait et les arches ont été remplacés par des boucles et un lemniscate, formes ne présentant aucun point de rebroussement. Les quatre formes sont présentées sur la figure 2.5(a).

Sons enregistrés.  Les stimuli correspondant aux cercles et aux ellipses sont les mêmes que dans l'expérience 2. Pour les boucles et les arches, les sessions d'enregistrement se sont déroulées dans les mêmes conditions que celles de l'expérience 2. De même, les séquences correspondant à quatre périodes des sons enregistrés ont été sélectionnées sur la base des caractéristiques géométriques et temporelles du scripteur, et sont décrites dans la table 2.3. La figure 2.5(b) présente les profils de

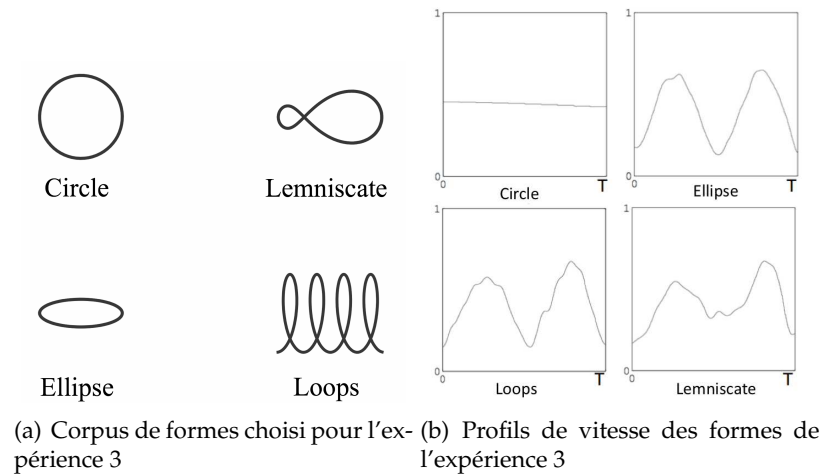


FIGURE 2.5 – Formes et profils de vitesse associés de l'expérience 3 – Figure issue de Thoret *et al.* (2014c).

vitesse sur une période des quatre formes. Il est important de noter que ces quatre profils sont plus proches les uns des autres que ceux de l'expérience 2. Si notre hypothèse est vérifiée, cela devrait impliquer plus de confusions perceptives entre les sons. Seul le cercle semble avoir toujours un profil de vitesse réellement distinct des trois autres formes.


Sons de synthèse.  Les sons de frottement synthétiques ont été générés à partir des profils de vitesse enregistrés à l'aide de la tablette graphique en utilisant le même modèle de synthèse que dans les expériences 1 et 2.

TABLE 2.3 – Caractéristiques géométriques et temporelles des séquences choisies pour l'expérience 3

Forme	Longueur théorique (cm)	Longueur (cm)	Durée (secondes)
Cercle	62.02	62.5	5.2
Ellipse	92.04	89.32	5.8
Lemniscate	147.48	145.46	5.6
Boucles	90.73	92.1	5.4

2.5.2 Résultats

Les résultats, présentés en table 2.4, révèlent que les sons ont été concrètement associés aux formes avec des taux de réussite significativement supérieurs à la chance à la fois pour les sons de synthèse et les sons enregistrés, exceptés dans le cas des boucles pour les sons enregistrés avec un taux de réussite de seulement 29.17% ($p < .001$ pour chaque forme excepté pour les boucles, $d_{\text{cercle}} = 6.12$; $d_{\text{ellipse}} = 0.53$; $d_{\text{lemniscate}} = 1.26$; $d_{\text{loops}} = 0.13$ pour les sons enregistrés, $d_{\text{cercle}} > 100$; $d_{\text{ellipse}} = 1.30$; $d_{\text{lemniscate}} = 2.23$; $d_{\text{boucle}} = 0.7$ pour les sons de synthèse). Le cercle et le lemniscate présentent les plus hauts taux de réussite et aucune confusion avec les autres formes n'a été observée.

A l'inverse, les ellipses et les boucles, bien que leurs taux de réussite soient au dessus de la chance statistique, ont été confondues à la fois pour les sons enregistrés (51.39% et 45.83%) et pour les sons de synthèse (45.83% et 43.06%). Ces taux d'association ne diffèrent pas significativement les uns des autres que ce soit pour les sons enregistrés, $t(17) = 0.73$; $p = .47$ et $t(17) = 1.89$ et $p = .22$, ou pour les sons de synthèse, $t(17) = 0.77$; $p = .66$ et $t(17) = 0$; $p = 1$.

Les résultats révèlent également que les matrices d'associations des résultats pour les sons de synthèse et des sons enregistrés sont significativement corrélées, $r(4) = .94$, $p < .001$. De plus, pour chaque forme les tests de Wilcoxon montrent que les taux de réussite ne diffèrent pas entre les deux types de sons (Circle : $z = -1$, $p = 1$; Ellipse : $z = -.99$, $p = .32$; Lemniscate : $z = -1.348$, $p = .18$; Boucle : $z = -1.29$, $p = .20$). Cela révèle, comme pour l'expérience 2, que les deux types de sons ont procuré des taux d'associations similaires.

TABLE 2.4 – Matrices d'association sons/formes – Moyennes et erreurs standards en pourcentage – Expérience 3. Les taux d'association supérieurs à la chance (25%) sont notés en gras. Le niveau de significativité des comparaisons à la chance est noté : * $p < .05$, ** $p < .01$, *** $p < .001$

Forme	Sons enregistrés			
	Cercle	Ellipse	Lemniscate	Boucles
Cercle	97.22***	2.78	-	-
SE	3.9	3.9	-	-
Ellipse	-	41.67***	6.94	51.3***
SE	-	11.7	6.03	11.86
Lemniscate	2.78	9.72	68.06***	19.4
SE	3.9	7.03	11.07	9.39
Boucles	-	45.83***	25	29.17
SE	-	11.83	10.28	10.79
Forme	Sons synthétisés			
	Cercle	Ellipse	Lemniscate	Boucles
Cercle	100***	-	-	-
SE	0	-	-	-
Ellipse	-	50***	4.17	45.83***
SE	-	11.87	4.74	11.83
Lemniscate	-	6.94	81.94***	11.1
SE	-	6.03	8.55	7.46
Boucles	-	43.06**	13.89	43.06**
SE	-	11.75	8.21	11.75

2.5.3 Discussion

Les résultats de cette expérience confirment ceux de l'expérience 2, le cercle a toujours été parfaitement associé au bon son à la fois pour les sons de synthèse et les sons enregistrés bien que les trois autres sons se discriminent moins aisément du point de vue perceptif. Cependant, les taux de réussite pour l'ellipse sont inférieurs à ceux de l'expérience 2. De plus, les résultats révèlent des confusions entre les boucles et l'ellipse. Le lemnicate a été bien identifié dans l'ensemble, malgré quelques confusions avec les boucles et l'ellipse. Ces confusions peuvent s'expliquer par la proximité des profils de vitesse de ces différentes formes, voir figure 2.5(b). Finalement, aucune différence

entre les résultats obtenus pour les sons enregistrés et les sons synthétisés à partir du profil de vitesse n'a été mise en évidence. Cependant, bien que les différences ne soient pas significatives, on peut noter que les sons de synthèse ont procuré des taux de réussite plus grands que les sons enregistrés, en particulier dans le cas de l'ellipse pour laquelle la différence entre les taux de réussite est supérieure à 10%.

Ces résultats renforcent ceux de l'expérience 2 concernant la pertinence du profil de vitesse dans l'intégration auditive du geste sous-jacent à un son de frottement. Cependant, les confusions observées entre les ellipses et les boucles dans l'expérience 3 montrent que cette information n'est pas suffisante pour discriminer l'intégralité de n'importe quelle forme géométrique, et en particulier pour discriminer les formes avec des géométries proches. Cette confusion va dans le sens de la loi en puissance 1/3 qui stipule que la vitesse du geste est contrainte par la courbure de la forme dessinée. Ainsi, puisque les ellipses et les boucles ont des géométries proches et donc des courbures similaires, les profils temporels de vitesse générés en les dessinant ne se distinguent pas suffisamment pour reconnaître ces formes à travers les sons.

2.6 Le profil de vitesse : un invariant perceptif

L'analyse supplémentaire présentée dans cette section est extraite du chapitre d'ouvrage :

Thoret, E., Aramaki, M., Kronland-Martinet, R., Velay, J. L., & Ystad, S. (2013). Reenacting sensorimotor features of drawing movements from friction sounds. In *From Sounds to Music and Emotions*, LNCS vol 7900, 130-153. Springer Berlin Heidelberg

Les résultats présentés dans les expériences 2 et 3 ont montré que les sons de synthèse sont aussi efficaces que les sons enregistrés pour aider à la reconnaissance des formes simples, suggérant ainsi que le profil de vitesse reflété dans leurs variations de timbre était l'information utilisée par les sujets pour effectuer l'association entre le son et la forme. Afin d'appuyer ces résultats perceptifs, cette section présente une comparaison entre les classifications des formes obtenues à partir des résultats perceptifs et à partir d'une comparaison *mathématique* des profils de vitesse. L'objectif de cette analyse supplémentaire est de montrer que les résultats des expériences étaient prévisibles à partir de l'analyse des profils de vitesse, ce qui validerait ainsi notre hypothèse initiale sur la pertinence du point de vue cognitif profil de vitesse pour la reconnaissance du geste par le son.

2.6.1 Classification objective des formes en fonction des profils de vitesse

En pratique, les profils de vitesse enregistrés v_i sont échantillonnés à une fréquence de 200 Hz. Un profil de vitesse correspond donc à un vecteur de N points dont la taille varie selon la durée de l'enregistrement. Les durées des profils des différentes formes n'étant pas les mêmes pour chaque forme, il est donc nécessaire de pouvoir comparer deux vecteurs de tailles différentes pour associer deux profils de vitesses v et w .

Distance euclidienne entre deux profils de vitesse.

Pour calculer une distance entre deux profils il est nécessaire de définir une métrique. Le choix de la métrique est crucial car il définit la façon dont la distance entre les formes va être calculée. La métrique la plus utilisée est la métrique euclidienne qui correspond au produit scalaire suivant entre

deux vecteurs v_i et w_i de même longueur :

$$\langle v|w \rangle = \sum_{k=1}^N v_k w_k \quad (2.4)$$

La distance euclidienne d entre les deux profils est alors définie par $d(v, w) = \|v - w\| = \sqrt{\langle v - w | v - w \rangle}$. Cette distance est minimale pour $v = w$ et augmente à mesure que la différence entre v et w augmente. Dans le cas des profils de vitesse enregistrés, afin de pouvoir se ramener à des vecteurs de taille égale, un rééchantillonnage sur 512 points et une normalisation en fonction de leurs énergie ont été effectués. Cela se justifie aisément dans la mesure où les durées des sons sont très proches, cf. tables 2.1 et 2.3. Le rééchantillonnage sur un nombre de pas constant ne modifie donc pas significativement leurs profils et les dilatations temporelles appliquées sont très faibles. Des algorithmes plus complexes, tels que le Dynamic Time Warping (Keogh et Ratanamahatana, 2005) existent afin de calculer une distance entre deux vecteurs de tailles différentes, cependant ces algorithmes ne procurent pas d'avantage significatif dans le cas de notre étude.

Matrice de distance et clustering.

La distance euclidienne entre deux profils de vitesse permet de calculer une matrice de distance entre chaque paire de profil pour chaque corpus de sons (6 paires pour chaque corpus). Les valeurs de la diagonale de cette matrice sont nulles, en effet la distance entre un profil et lui même est nulle. Pour chaque corpus, un dendrogramme ascendant (avec complete linkage) correspondant aux matrices de distance a ainsi été calculé voir figure 2.6. Pour les deux corpus, les dendrogrammes confirment les observations empiriques, le cercle est la forme la plus différente des autres quel que soit le corpus. Cela s'explique aisément par la forme très particulière de son profil de vitesse constant. Dans le premier corpus, on peut remarquer que les arches sont à peu près équidistants de l'ellipse et du trait. Dans le second corpus, l'ellipse et les boucles sont bien les formes qui sont les plus proches du point de vue des profils de vitesse.

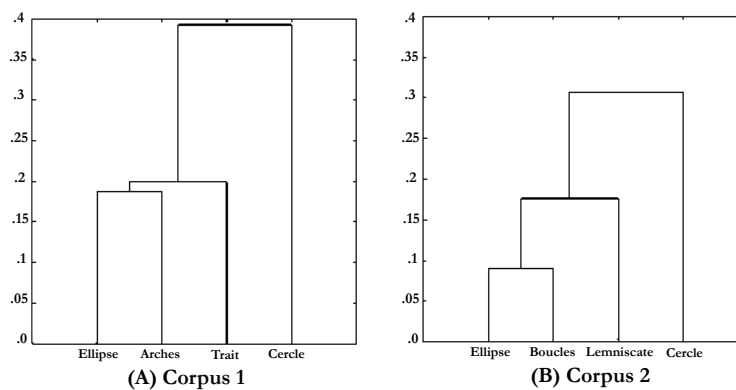


FIGURE 2.6 – Panels A et B : Dendrogrammes ascendants hiérarchiques calculés à partir de la matrice de distance pour les corpus 1 et 2 respectivement – Figure issue de Thoret *et al.* (b).

2.6.2 Comparaison entre classifications perceptives et mathématiques

A partir des matrices d'association S des tests perceptifs des expériences 2 et 3, une matrice de distances $D = 1 - S^*$ a été calculée à partir d'une version symétrisée S^* de S :

$$S^* = \frac{S + S^t}{2} \quad (2.5)$$

Ainsi, pour chaque expérience, i.e. chaque corpus, deux matrices de distances ont été calculées, une pour les sons enregistrés et une pour les sons synthétisés. De la même façon que pour les profils de vitesse dans le paragraphe précédent, un dendrogramme a été calculé à partir de ces matrices de distance. Finalement, quatre dendrogrammes (présentés en figure 2.7) ont été calculés à partir des résultats des expériences 2 et 3. Une observation globale des quatre classifications permet de supposer que les sons de synthèse et les sons enregistrés ont bien procuré des classifications perceptives équivalentes. En effet, pour chaque corpus, la proximité relative entre chaque forme semble avoir été conservée. Afin de valider statistiquement cette observation, les corrélations de Pearson et Spearman sur la base des distances cophénétiques ont été calculées pour chaque corpus entre les classifications issues des résultats obtenus avec les sons de synthèse et celles issues des résultats obtenus à partir des sons enregistrés. Les comparaisons entre les dendrogrammes obtenus pour les sons de synthèse

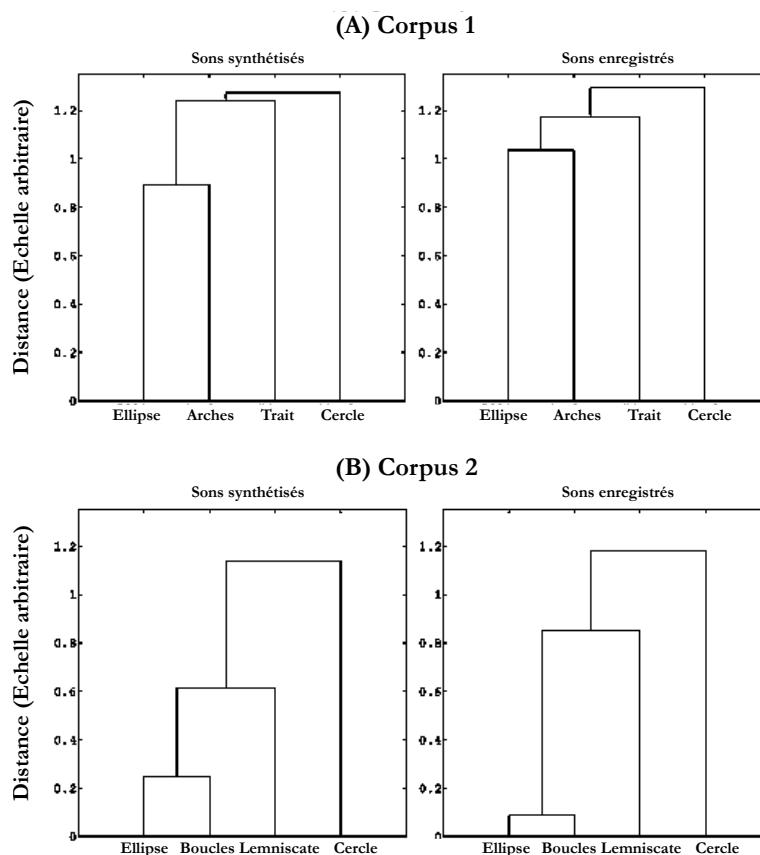


FIGURE 2.7 – Panel A : Clusterings hiérarchiques ascendants calculés à partir des matrices de confusion de l’expérience 1 - Des corrélations significatives ont été trouvées entre les distances cophénétiques des deux dendrogrammes ($r = .89$ and $\rho = 1.$) — Panel B : Clusterings hiérarchiques ascendants calculés à partir des matrices de confusion de l’expérience 2 - Des corrélations significatives ont été trouvées entre les distances cophénétiques des deux dendrogrammes ($r = .89$ and $\rho = 1.$) – Figure issue de Thoret *et al.* (b).

et sons enregistrés ont déjà été présentées et ne présentent aucune différence significative, cf. paragraphes 2.4.2 et 2.5.2. Les comparaisons entre ces dendrogrammes et les classifications objectives obtenues à partir des distances euclidiennes entre profils de vitesse révèlent qu’elles ne diffèrent pas significativement entre elles, excepté pour les sons enregistrés dans le cas de l’expérience 2 (formes distinctes) où la corrélation $r = .58$ n’est pas significative. La table 2.5 présente les coefficients de corrélation.

TABLE 2.5 – Coefficients de corrélation de Pearson et Spearman, notés respectivement r et ρ , entre les distances cophénétiques des classifications perceptives et mathématiques — Les corrélations significatives sont en gras.

	Expérience 2		Expérience 3	
	r	ρ	r	ρ
Enregistrés	.58	.93	.98	.93
Synthétisés	.89	.93	.94	.93

2.6.3 Discussion

Les classifications présentées dans cette analyse confirment donc que le profil de vitesse du geste est bien l'information perceptive utilisée dans la reconnaissance auditive d'un mouvement humain et des formes géométriques. De fait, il peut être considéré comme un invariant transformationnel permettant la reconnaissance sonore des gestes humains, et même de discriminer des formes géométriques. En effet, les comparaisons entre classifications objectives, basées sur les profils de vitesse, et les classifications obtenues à partir des résultats perceptifs des expériences 2 et 3 sont significativement corrélées.

2.7 Discussion générale

La série d'expériences présentée dans cette étude s'est intéressée à la capacité des êtres humains à percevoir les mouvements biologiques à travers les sons, et même à la possibilité de discriminer des formes géométriques simples sur la bases des sons. Cette étude est la première à notre connaissance qui met en évidence de façon aussi formelle la relation entre la perception auditive et les mouvements biologiques. Dans l'expérience 1, en s'appuyant sur un protocole mis en place par Viviani et Stucchi (1992) pour la perception visuelle, il a été mis en évidence que le profil de vitesse des gestes humains était un indice dynamique pertinent pour la reconnaissance des gestes. Afin de mettre en place le protocole, un modèle de synthèse sonore de son de frottements a été utilisé pour générer des variations de timbre en accord avec le profil de vitesse du mouvement. L'utilisation de la synthèse sonore a permis de se focaliser uniquement sur ce paramètre, et ainsi de révéler que pour évoquer un mouvement humain fluide ou naturel, les variations de timbre doivent respecter la loi en puissance $1/3$, signature cinématique des mouvements biologiques.

Les expériences 2 et 3 se sont elles attachées à évaluer dans quelle mesure il était possible de déterminer les caractéristiques spatiales d'une forme géométrique à partir d'un son dans une tâche d'association entre les sons de frottement produits par une personne qui dessine et des formes géométriques statiques. Des sons enregistrés et des sons de synthèse ont été utilisés permettant ainsi d'évaluer l'importance du profil de vitesse dans la perception auditive des sons produits par les mouvements biologiques. Dans l'expérience 2, un corpus de formes aisément différenciables du point de vue géométrique et sonore a été réalisé. Dans l'expérience 3, un corpus de formes plus proches a été réalisé. Les taux de réussite ont été comme prévu élevés dans le cas des formes distinctes et des confusions sont apparues dans le cas des formes plus proches. Discriminer des formes géométriques sur la base du son est donc possible dans la mesure où les sons sont assez différents. De plus, sur la base de comparaisons entre les matrices d'associations, il a été mis en évidence que les sons de synthèse avaient des taux d'associations similaires aux sons enregistrés, suggérant ainsi que le profil de vitesse contenu dans les variations de timbre des sons de synthèse était l'information utilisée par les

sujets pour réaliser l'association entre le son et la forme. De plus, cette hypothèse a été appuyée par la comparaison avec une classification objective des profils de vitesse qui a fourni une classification comparable à celle des résultats perceptifs.

Dans une perspective écologique, les variations spécifiques de timbre inhérentes au profil de vitesse des gestes humains peuvent donc être considérées comme directement reliées à un invariant transformationnel permettant la reconnaissance sonore d'un mouvement. Ces variations de timbre spécifiques affordent l'action de dessiner de façon fluide et naturelle, et ce résultat est à mettre en regard des études sur la perception auditive des événements sonores (Gaver, 1993b; Li *et al.*, 1991; Repp, 1987; Warren et Verbrugge, 1984). En complément de cette analyse relevant de la perspective écologique, les résultats précédents peuvent également être discutés dans le cadre du contexte idéo-moteur (Hommel *et al.*, 2001; Prinz, 1997). Dans une étude récente, Young *et al.* (2013) ont demandé à des sujets de reproduire en temps réel des patterns de bruits de pas, i.e. les sons produits par une personne en train de marcher, les patterns évoquaient des foulées de différentes longueurs. Ils ont ainsi pu observer que les sujets reproduisaient de façon précise les caractéristiques des pas évoqués par les sons. De plus, de la même façon que ce que nous avons mis en évidence avec les sons de synthèse, les sujets étaient capables d'effectuer cette tâche même quand l'information cinématique seule était présente dans des bruits de pas synthétisés. Cela suggère que la perception auditive d'une action semble activer les mêmes schémas moteurs que quand l'action est réalisée, et que finalement, écouter une action permet de s'imaginer les caractéristiques de l'action et même de la reproduire. Afin d'expliquer cette relation audio-motrice, les auteurs ont suggéré que la perception et la planification du mouvement impliquaient les mêmes représentations cognitives. Cette description en terme de représentations cognitives, et donc *a priori* opposée à l'approche écologique, a été suggérée dans des études récentes tentant de réconcilier les éléments théoriques des approches directes, i.e. écologiques, et indirectes, i.e. cognitivistes, de la perception et de l'action (Cisek et Kalaska, 2010; Norman, 2002). Dans le cas de leurs expériences, Young *et al.* (2013) ont fait l'hypothèse que la tâche avait été correctement réalisée grâce à l'existence d'un percept audio-moteur commun et unifié. Dans le cadre de nos expériences, nous suggérons l'existence d'un même percept audio-moteur pour la perception sonore des sons de frottement produits par les mouvements graphiques. Et en particulier le fait que, inférer une forme graphique par le biais d'un son nécessite une association interne de ces deux percepts. Cette relation est loin d'être triviale car *a priori* aucun son n'est naturellement associé à une forme géométrique. Cependant, les résultats de l'expérience 1 et les études présentées en introduction mettent en évidence que percevoir une forme géométrique, aussi bien du point de vue visuel que sonore, implique que l'on s' imagine en train de la dessiner et que l'on s' imagine donc indirectement le profil de vitesse du geste que nous réaliserions pour la produire. Cela implique donc que nous traitons l'information cinématique des gestes à un niveau amodal ne relevant *a priori* pas d'une modalité en particulier, comme ce qui avait été suggéré par Viviani *et al.* (1997). En effet, nos expériences ne permettent pas de déterminer si les sujets ont réalisé les associations sons/formes en déduisant la forme depuis le son ou bien le son depuis la forme. La représentation sensori-motrice des mouvements biologiques constitue donc la pierre angulaire de ce processus d'association inter-modal. Ce résultat renforce de plus celui de Viviani et Stucchi (1989) qui ont mis en évidence que la relation entre la représentation dynamique d'une forme, i.e. un point lumineux en mouvement, et une forme statique, était possible par le biais d'une représentation sensori-motrice du mouvement produit pour dessiner la forme.

Ce premier chapitre a donc mis en évidence l'importance du profil de vitesse et cela a été rendu possible grâce à l'utilisation de sons de frottement générés uniquement à partir du profil de vitesse.

On voit ici tout l'intérêt du paradigme action → objet pour générer des stimuli calibrés permettant de n'évoquer qu'un seul paramètre du geste et ainsi d'en évaluer l'importance perceptive.

Du point de vue de la synthèse sonore et du contrôle des modèles de synthèse, ces résultats sont également très intéressants car ils permettent d'envisager la création de contrôles sémantiques haut-niveau reliés à la nature du geste, i.e. fluide/naturel, et par rapport à des formes géométriques. Il est ainsi envisageable d'approfondir les travaux d'Adrien Merer (2011) sur la création de synthétiseurs de sons évoquant des mouvements en y ajoutant des contrôles spécifiques reliés à un mouvement humain, tout cela dans le but de générer des sons évoquant des mouvements humains en les décrivant de façon très simple de façon sémantique ou même par des gestes.

Chapitre 3

Ecouter les sons produits par des mouvements biologiques révèle les prototypes géométriques

Sommaire

3.1	Introduction	71
3.2	Expérience 4 - Prototype visuel statique	76
3.3	Expérience 5 - Prototype moteur	80
3.4	Expérience 6 - Prototype visuel dynamique	83
3.5	Expérience 7 - Prototype sonore	87
3.6	Comparaison entre les différents prototypes	90
3.7	Expérience 8 - Comparaison entre les métriques sonore et visuelle	93
3.8	Discussion générale	98

Dans le chapitre 2, nous avons donc mis en évidence que le profil de vitesse des gestes humains était un invariant transformationnel de la reconnaissance auditive des gestes graphiques. En utilisant des sons de frottement de synthèse, nous avons en effet montré que nous étions capable de reconnaître la loi en puissance $1/3$, reliant la vitesse tangentielle d'un geste à la courbure de la forme dessinée. Dans deux autres expériences nous avons montré que nous étions capables, dans une certaine mesure de discriminer des formes géométriques juste sur la base des sons de frottement produits par une personne qui dessine. Enfin, en comparant les résultats obtenus dans des tâches de discrimination faites avec des sons de synthèse générés uniquement à partir du profil de vitesse, nous avons montré que cette discrimination était basée sur une comparaison perceptive entre les profils de vitesse des gestes qui ont produits les sons de frottement.

Les résultats de ces deux dernières expériences suggèrent donc que l'on est capable d'associer une forme au son de frottement produit par le biais du geste évoqué. Cependant, nous avons seulement montré que l'on pouvait discriminer ces formes, nous ne savons absolument pas quelle est la géométrie exacte que le sujet s'est réellement imaginée à partir du son. Par exemple la géométrie de l'ellipse imaginée est-elle cohérente avec celle qui était affichée sur l'interface du test ? Rien ne certifie cela et les sujets ont très bien pu imaginer une ellipse plus plate ou plus arrondie que celle qui était affichée. Dans ce chapitre nous avons donc souhaité mieux comprendre la relation entre le son et la

forme imaginée par le sujet. Nous allons nous focaliser ainsi sur le cas particulier des ellipses, qui sont des formes intermédiaires entre un trait et un cercle.

3.1 Introduction

Quelle image le peintre a-t-il dans la tête avant de poser le pinceau sur la toile et de la traduire gestuellement ? Quel son le compositeur imagine-t-il avant de poser les mains sur le piano ? La façon dont notre cerveau s' imagine les objets et les actions qu'il réalise est une question qui a été abordée depuis de nombreux siècles. De l'allégorie de la caverne de Platon, qui suggérait que la mémoire se fonde sur des images du monde réel, jusqu'aux recherches actuelles en psychologie cognitive et neurosciences qui observent le cerveau en train d'imaginer et raisonner en temps réel, il a été mis en évidence que les caractéristiques des images mentales qui apparaissent à notre esprit par anticipation quand on imagine un objet, une personne, ou encore une action, sont reliées aux caractéristiques réelles de ces objets ou de ces actions (Shepard et Metzler, 1971; Shepard et Judd, 1976; Jeannerod, 1995). Par ailleurs, notre perception du monde étant multimodale par essence, ces images se forment donc par rapport à l'ensemble des informations, conséquences de nos actions et de nos mouvements, qui arrivent de façon synchrone à nos différents canaux sensoriels. Ainsi, une action aussi simple qu'imaginer une forme géométrique comme un cercle requiert d'avoir déjà au moins une fois vu cette forme, ou bien de l'avoir produite avec ses mouvements (dans le cas d'un aveugle de naissance par exemple). Les caractéristiques de cette image peuvent donc aussi bien relever de notre expérience visuelle que de notre expérience motrice.

De façon plus surprenante, nous avons mis en évidence que les formes géométriques peuvent également être transmises par le biais de la modalité sonore. On a vu dans le chapitre précédent que les sons de frottement produits par une personne qui dessine contiennent assez d'informations sur le geste qui a dessiné une forme, et plus précisément sur sa vitesse, pour discriminer des formes simples uniquement en écoutant ces traces sonores du geste. De plus, une expérience a montré que des sujets étaient capables de calibrer le profil de vitesse d'un geste afin que ses caractéristiques correspondent à celles des mouvements biologiques, i.e. la loi en puissance $1/3$. Cela signifie que la façon dont on s' imagine les formes est intrinsèquement reliée à la façon dont on les dessine. Ces expériences ont montré qu'une telle association était possible dans le cadre d'une tâche de discrimination, cependant, elles ne permettent pas de connaître la forme réelle que l'on s' imagine quand on entend le son produit par une personne qui dessine.

Afin de révéler la forme qu'une personne s' imagine en entendant le son produit par quelqu'un qui dessine, il est nécessaire de s'intéresser à la façon dont on s' imagine ces formes au niveau visuel. Chaque type de forme géométrique, un cercle, une ellipse, un rectangle... peut être considéré comme un ensemble intégrant un nombre infini d'éléments particuliers qui partagent des propriétés communes (Piaget et Matalon, 1958; Theurel *et al.*, 2012). De nombreuses expériences ont mis en évidence que du point de vue cognitif, certaines de ces formes ont un statut particulier au sein d'une catégorie et qu'elles sont les plus représentatives de toutes les formes de la catégorie dans leur ensemble. Ces formes particulières, également appelées formes prototypiques, ont été mises en évidence aussi bien dans des tâches perceptives que dans des tâches de production motrice. Il existe ainsi différentes manières de les mettre en évidence. Du point de vue perceptif, des expériences de jugement de typicité (Rosch, 1973; Rosch et Mervis, 1975; Feldman, 2000), d'apprentissage (Rosch, 1973, 1975), de reconnaissance (Theurel *et al.*, 2012) ou encore de discrimination (Wamain *et al.*, 2011) ont été réalisées. Du point de vue de la production motrice, lors de la plupart des expériences qui ont été réalisées, il était demandé aux sujets de dessiner une forme le plus spontanément possible

(Kalénine *et al.*, 2013) ou de dessiner plein d'exemplaires différents d'une même forme (Feldman, 2000) afin de déterminer quelle était la forme qui était la plus caractéristique de cette catégorie du point de vue moteur. Dans toutes ces études, il a été montré que nous catégorisons les formes autour de prototypes géométriques. Feldman (2000) s'est par exemple intéressé à comparer les prototypes perceptifs et moteurs pour les quadrilatères et les triangles de façon très générale. Ils ont mis en évidence que les formes les plus représentatives de ces catégories étaient les carrés et triangles équilatéraux. Kalénine *et al.* (2013) ont eux montré que des enfants et des adultes dessinaient des rectangles et des triangles dans des proportions similaires centrées autour d'une certaine valeur et ayant une orientation préférentielle, cependant les prototypes perceptifs (visuels) différaient des prototypes moteurs. Les rectangles dessinés avaient en moyenne tous un ratio longueur sur largeur de l'ordre de 2.3 et étaient tous dessinés avec la longueur orientée à l'horizontale. Theurel *et al.* (2012) ont eux montré l'existence de prototypes haptiques dans une tâche de reconnaissance tactile de formes géométriques. Au-delà de la mise en évidence des caractéristiques de chacun des prototypes, un des enjeux principaux de toutes ces études est de comprendre l'origine de leur géométrie particulière. Différentes hypothèses ont d'ores et déjà été formulées et évaluées. Celle qui semble faire consensus est que les prototypes émanent principalement d'une expérience visuelle. Theurel *et al.* (2012) ont en effet montré qu'il existait un prototype haptique du carré, du rectangle, et du triangle chez des sujets voyants qui avaient les yeux bandés mais ils ont constaté que ce prototype n'existait pas chez des sujets aveugles de naissance, n'ayant ainsi eu aucune expérience visuelle. La modalité visuelle semble donc être la condition nécessaire à l'émergence d'un prototype. Cependant, de nombreuses théories et expériences ont également montré l'importance de l'interaction sensori-motrice dans l'émergence des concepts sémantiques tels que les formes géométriques (Barsalou, 1999).

Les formes prototypiques semblent donc avoir une position très particulière dans les processus de catégorisation et de représentation des formes géométriques. Par conséquent, on peut faire l'hypothèse qu'elles ont également une position importante quand un sujet va s'imaginer une forme à partir du son de frottement produit par les gestes d'un scripteur. Cette association son/forme par le biais du geste évoqué positionne le geste évoqué au milieu de ce processus d'association. Au regard des connaissances précédentes sur la notion de forme prototypique émanant principalement d'une expérience visuelle, plusieurs questions importantes doivent être posées. Existe-t-il un prototype dynamique des formes géométriques révélé par les sons de frottement produits quand une personne dessine ? Si oui, ce prototype est-il commun avec les prototypes visuels et moteurs ? Ces questions sont cruciales car leur donner une réponse permettrait d'apporter de nouvelles hypothèses sur l'origine de l'émergence de prototypes géométriques, et en particulier du rôle de la motricité et de la boucle sensori-motrice dans la caractérisation de leurs géométrie. D'autre part, répondre à ces questions permettra de paramétrer de façon précise des processus de sonification de mouvements ou de formes de façon à ce que les sons synthétisés aient des caractéristiques acoustiques qui évoquent les bonnes trajectoires.

Afin d'aborder ces problèmes nous allons nous intéresser à un type de forme particulier, les ellipses. Les formes présentant des points de rebroussement comme les carrés, les rectangles ou les triangles, génèrent des mouvements avec des vitesses nulles et donc des sons de frottement avec des silences, indices acoustiques très saillants et pouvant captiver l'attention. Les formes dont la courbure ne s'annule jamais semblent donc dans un premier temps plus adaptées pour étudier les questions soulevées dans le paragraphe précédent. De plus, les ellipses sont des formes qui ont été largement étudiées, aussi bien du point de vue de leur production que de leur perception visuelle. Du point de vue moteur, ces formes ont très largement été étudiées et en particulier, il a été mis en

évidence que les caractéristiques dynamiques et géométriques de leurs productions pouvaient être représentées par des modèles très simples.

Les mécanismes biomécaniques de production des mouvements graphiques (comme l'écriture) ont en effet été largement étudiés depuis très longtemps. L'approche dynamique de la production des mouvements soutient que l'on peut les modéliser comme le résultat du mouvement d'oscillateurs harmoniques de fréquences et amplitudes évoluant au cours du temps (Hollerbach, 1981). Ils proposent que le mouvement $(x(t), y(t))$ de la main d'une personne qui dessine dans le plan est régi par les équations suivantes :

$$\begin{cases} x(t) = A_x \cos(\omega_x t + \phi_x) \\ y(t) = A_y \cos(\omega_y t + \phi_y) \end{cases} \quad (3.1)$$

où A_x et A_y sont les amplitudes, ω_x et ω_y les fréquences, et ϕ_x et ϕ_y les phases des deux oscillateurs. Haken *et al.* (1985) ont montré que ce modèle pouvait se dériver d'une modélisation mécanique de la coordination dynamique des mouvements mettant en jeu des oscillateurs harmoniques couplés non linéairement. Ce type de système a des caractéristiques phénoménologiques comparables à celles des systèmes auto-organisés mettant en évidence l'existence d'états stables (Kelso *et al.*, 1981; Kelso, 1984, 2012), aussi appelés attracteurs moteurs. En effet, le nombre de degrés de liberté en jeu dans la coordination des mouvements est tellement grand qu'il paraît improbable que le système nerveux central puisse les contrôler tous en même temps (Danna, 2011). L'évolution semble avoir *trouvé*¹ ici une solution fonctionnelle permettant d'atteindre des états en ne contrôlant que très peu de paramètres. Ainsi, les transitions entre ces différents états stables sont contrôlées par la modification de paramètres de très haut niveau comme le bruit injecté dans le système. Dans le cas de mouvements circulaires ou elliptiques, on peut faire les approximations suivantes : $A_x = A_y$ et $\omega_x = \omega_y$ permettant de simplifier le modèle de la façon suivante :

$$\begin{cases} x(t) = A \cos(\omega t) \\ y(t) = A \cos(\omega t + \phi) \end{cases} \quad (3.2)$$

où A est l'amplitude de l'ellipse, ω la fréquence du mouvement, et $\phi = \phi_x - \phi_y$, la phase relative entre les deux oscillateurs. La phase relative ϕ détermine la géométrie de la trajectoire elliptique parcourue (voir figure 3.1), en particulier, ϕ peut être reliée à l'excentricité de l'ellipse parcourue. Quand $\phi = 0$ (i.e. 0 degrés), les deux oscillateurs sont en phase et la trajectoire correspond à un

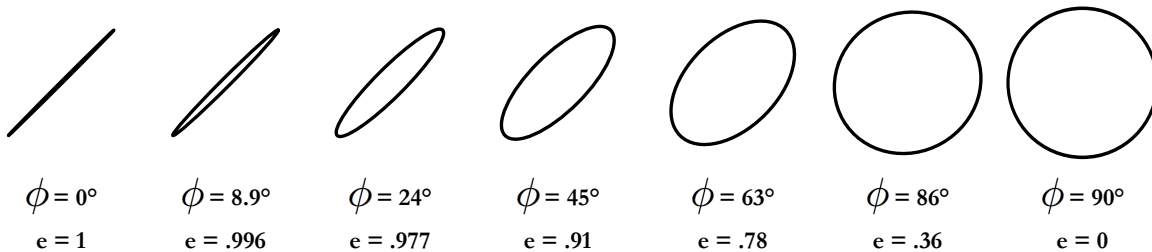


FIGURE 3.1 – Exemples de l'effet de la variation de la phase relative ϕ sur l'excentricité e de l'ellipse

trait, quand $\phi = \frac{\pi}{2}$ (i.e. 90 degrés), les deux oscillateurs sont en quadrature de phase et la trajectoire parcourue est un cercle. Toutes les valeurs de ϕ comprises entre 0 et 90 degrés correspondent à des ellipses intermédiaires entre le trait et le cercle. En pratique, l'excentricité et la phase relative ϕ sont

1. Cette perspective finaliste de l'évolution est ici employée dans un seul but didactique, l'évolution des êtres vivants étant le résultat de mutations, sélections et propagations de caractères avantageux au sein d'un environnement particulier, il apparaît inconcevable de considérer la moindre intervention d'un agent cognitif derrière ces évolutions.

reliées par les relations suivantes² :

$$\begin{cases} \phi = 2 \arctan \sqrt{1 - e^2} \\ e = \sqrt{1 - \tan^2 \frac{\phi}{2}} \end{cases} \quad (3.3)$$

Un trait correspond donc à une ellipse d'excentricité 1, un cercle à une ellipse d'excentricité nulle, et toutes les autres ellipses ont une excentricité comprise entre 0 et 1, voir figure 3.1. Cette modélisation offre donc un regard très simple sur la façon dont les humains produisent ce type de forme et des études ont montré qu'elle est très bien adaptée pour décrire les phénomènes biomécaniques impliqués dans la production des mouvements. Un point qu'il est essentiel de noter à ce stade est que cette modélisation du mouvement par deux oscillateurs harmoniques déphasés respecte la loi en puissance 1/3 (Lacquaniti *et al.*, 1983) caractéristique des mouvements humains déjà introduites dans le chapitre précédent, une démonstration est présentée en annexe 7.1.

Dans le cas des formes closes comme les ellipses, Dounskaia *et al.* (2000); Athenes *et al.* (2004) ont montré que si l'on demande à une personne de dessiner de façon continue et plusieurs fois de suite très vite un cercle, la phase relative entre les deux oscillateurs va tendre vers une valeur proche de 45 degrés, correspondant à une ellipse d'excentricité 0.9. Dans une comparaison entre une expérience psychophysique et une expérience motrice, Wamain *et al.* (2011) ont montré que l'on discrimine visuellement mieux les ellipses statiques dont l'excentricité (i.e. de phase relative) correspond à celle que l'on dessine préférentiellement, et que ces ellipses ont en moyenne toutes une valeur d'excentricité proche de .9 (excentricité correspondant à une phase relative de 45 degrés). Ces expériences pointent donc le fait que l'ellipse d'excentricité .9 et le mouvement associé de phase relative 45 degrés semblent être caractéristiques des trajectoires elliptiques, et peuvent ainsi suggérer qu'ils seraient de bons candidats pour caractériser le prototype de l'ellipse.

La littérature sur les prototypes présentée précédemment a fait état de grandeurs géométriques caractéristiques comme le ratio longueur sur largeur pour les rectangles. L'excentricité d'une ellipse peut être assimilée à un tel ratio et on peut faire l'hypothèse que si une ellipse prototypique existe, elle pourra être caractérisée par son excentricité. Le modèle présenté précédemment permet donc d'aborder la question du rôle de la motricité dans l'émergence des prototypes de façon très intéressante. Avec ce modèle, on peut en effet contrôler la géométrie de l'ellipse tout en conservant un profil de vitesse respectant la loi en puissance 1/3. A partir de ce modèle il est possible d'envisager un protocole expérimental permettant de déterminer le prototype de l'ellipse dans des tâches différentes, aussi bien visuelle (statique et dynamique), motrice, que sonore. Dans les cas visuel dynamique et moteur, le mouvement perçu intègre à la fois des informations géométriques et dynamiques, l'expérience sonore n'évalue elle que la dynamique car les variations de timbre d'un son de frottement n'évoque que la dynamique du mouvement, et l'on peut ainsi s'affranchir d'indices spatiaux directement reliés à la géométrie de l'ellipse. L'expérience sonore est donc la pierre angulaire de la série d'expériences qui va être présentée dans cette étude. En effet, la vision comme la motricité font intervenir aussi bien la géométrie que la dynamique d'un mouvement, excepté dans le cas visuel statique, et il n'est donc pas possible de trancher sur le rôle de la dynamique, et donc de la motricité, dans l'émergence d'un potentiel prototype. Seule l'audition permet d'évaluer la prototypicité de la dynamique par les variations intrinsèques d'un son si la dynamique en soi révèle l'existence d'une ellipse prototypique. Les stimuli utilisés seront des sons de synthèse correspondant à tout le panel d'ellipses décrit par le modèle précédent en générant le profil de vitesse tangentielle du mouvement à partir

2. Une démonstration est proposée en annexe 7.1

du système d'équations 3.2 :

$$v_t(t) = A\omega\sqrt{\sin^2(\omega t) + \sin^2(\omega t + \phi)} \quad (3.4)$$

Ce profil de vitesse sera ensuite utilisé pour générer des sons de frottement de synthèse en utilisant un modèle similaire que celui utilisé dans le chapitre 2 (Van Den Doel *et al.*, 2001; Thoret *et al.*, 2014c).

Finalement dans cette étude, on souhaite vérifier l'hypothèse que le prototype de l'ellipse correspond bien à une ellipse d'excentricité .9 correspondant à l'ellipse que l'on dessine préférentiellement, et que, plus que la géométrie de la forme, c'est la dynamique du geste sous-jacent qui est prototypique et qui est à l'origine de la géométrie du prototype. Pour cela, cinq expériences ont été réalisées afin d'évaluer dans différentes situations si un prototype de l'ellipse existe, et si oui :

1. par quelle excentricité est-il caractérisé ?
2. est-il commun à travers différentes modalités ?
3. son excentricité correspond-t-elle à celle de l'ellipse que l'on dessine préférentiellement (i.e. 0.9) ?

Dans la plupart des expériences présentées dans la littérature, la représentation d'une catégorie de forme par le biais d'une forme prototypique a été évaluée principalement avec des expériences de discrimination, de reconnaissance et de production motrice. Ici, pour les tâches perceptives, on a choisi de mettre en place des protocoles d'ajustement lors desquels les sujets devaient régler le stimulus (visuel ou sonore) afin qu'il leur évoque la forme la plus représentative de la catégorie ellipse. La première expérience (nommée *visuelle statique* dans la suite) s'est focalisée sur l'existence d'un prototype visuel *statique*, i.e. une forme affichée dans son intégralité, les sujets devaient régler l'excentricité d'une forme statique. Dans la seconde expérience (nommée *motrice* dans la suite) qui était une tâche de production motrice, afin d'évaluer l'existence d'un prototype moteur, les sujets devaient dessiner l'ellipse qui leur était la plus représentative. Par prototype moteur on se réfère à l'ellipse la plus représentative du point de vue d'une production motrice. Dans une troisième expérience (nommée *visuelle dynamique* dans la suite), les sujets devaient ajuster visuellement le mouvement d'un point parcourant une trajectoire elliptique pour que le mouvement soit le plus elliptique possible afin de mettre en évidence si le prototype visuel n'est pas avant tout caractérisé par une dynamique motrice. Ce prototype sera par la suite nommé prototype visuel dynamique. Enfin dans la quatrième expérience (nommée *sonore* dans la suite), les sujets devaient ajuster le son de frottement d'une personne qui dessine une trajectoire elliptique afin qu'il leur évoque l'ellipse la plus représentative. Ce prototype sera par la suite appelé prototype sonore. Les sons de frottement utilisés ont été synthétisés avec le modèle de synthèse présenté dans le chapitre 1. Afin d'évaluer une possible distorsion entre métrique visuelle et sonore et pour évaluer la géométrie de l'ellipse imaginée à partir des sons de frottement, une dernière expérience a été réalisée lors de laquelle les sujets devaient calibrer l'ellipse visuelle par rapport au son de frottement. En plus de poser une question centrale sur l'origine des prototypes, cette étude permet de poser le problème de la relation entre le son de frottement et la forme évoquée de façon très formelle. En effet, la dernière expérience permettra d'évaluer dans quelle mesure la géométrie de l'ellipse que l'on s' imagine à travers le son (et le geste évoqué) est cohérente avec l'ellipse théorique utilisée pour générer le son, répondant ainsi à un problème posé en sonification.

TABLE 3.1 – Amplitudes en *cm* des ellipses utilisées dans le test

A_1	A_2	A_3
3 cm	5 cm	10 cm

3.2 Expérience 4 - Prototype visuel statique


Afin d'évaluer l'existence d'un prototype visuel statique de l'ellipse défini par rapport à une excentricité préférentielle nous avons mis en place un protocole d'ajustement. Des sujets ont ajusté l'excentricité d'une ellipse statique sur un écran afin que l'ellipse affichée leur évoque l'ellipse la plus représentative de cette catégorie de forme. Une ellipse étant décrite par son excentricité mais également par sa *taille*, ce facteur a été introduit afin d'évaluer si il influe sur l'existence du prototype, et le cas échéant, sur l'excentricité du prototype.

3.2.1 Méthodes

Participants

20 sujets (8 femmes) d'un âge moyen de 30.6 ans (SD = 12.82) ont pris part volontairement à ce test. Tous les sujets étaient naïfs au sujet de l'expérience et avaient une vision normale ou corrigée. Tous les sujets ont donné leur consentement écrit avant de prendre part au test.

Stimuli

 Les stimuli visuels utilisés dans cette expérience sont composés d'ellipses statiques blanches sur fond noir de différentes tailles dont l'excentricité est déterminée par le modèle de geste présenté précédemment. Comme ces ellipses sont statiques, la notion de temps n'a plus de signification et on peut reformuler le système 3.2 en fonction de l'abscisse curviligne s de la façon suivante :

$$\begin{cases} x(s) = A \cos(2\pi s) \\ y(s) = A \cos(2\pi s + \phi) \end{cases} \quad (3.5)$$

où s est comprise entre 0 et 1. L'excentricité de l'ellipse e peut donc être aisément contrôlée en utilisant la relation mise en évidence précédemment :

$$\phi = 2 \arctan \sqrt{1 - e^2} \quad (3.6)$$

Ainsi, l'excentricité de l'ellipse peut être modifiée en manipulant implicitement la phase relative ϕ . (cf. figure 3.1). Il est important de noter qu'une inclinaison fixe de 45 degrés par rapport à l'horizontale a été utilisée dans cette expérience à l'instar de l'orientation de l'ellipse préférentielle déjà mise au jour dans le cadre de la production des mouvements elliptiques (Danna *et al.*, 2011). Trois ellipses de tailles différentes (i.e. l'amplitude A utilisée dans le système 3.5) ont été utilisées dans ce test :

Appareillage

Les sujets étaient assis devant un écran DELL 1907fp de résolution 1280 x 1024 et de fréquence de rafraichissement 60 Hz. Les ellipses étaient affichées au centre de l'écran. L'interface de test était programmée avec le logiciel Max/MSP.

Tâche

Il était demandé aux sujets d'ajuster l'excentricité de l'ellipse visuelle statique afin de régler l'ellipse qui leur paraissait la plus représentative de cette catégorie de formes géométriques. Le proto-

cole d'ajustement utilisé est basé sur un protocole proposé par Carlyon *et al.* (2010). Il était explicité ce à quoi correspondait la notion d'ellipse la plus représentative. Il leur était par exemple précisé qu'une ellipse était une forme entre un trait et un cercle, mais que lorsqu'ils imaginaient une ellipse, il pensait peut-être plus spécifiquement à une ellipse en particulier, et que c'était cette ellipse qu'ils allaient devoir régler sur l'écran. Pour régler l'excentricité, les sujets disposaient d'un clavier MIDI AKAI MPK avec 6 types de touches différentes : "<<<"; "<<"; "<" et ">"; ">>"; ">>>". En fonction de la touche, l'excentricité de l'ellipse était modifiée de façon plus ou moins importante, plus il y avait de flèches plus la modification était grande. Afin d'éviter un possible biais non-sensoriel, l'action des flèches, (i.e. augmenter ou diminuer l'excentricité), les flèches *de droite* et celles *de gauche* n'influaient pas de la même façon sur l'excentricité. Toutes les flèches allant dans la même direction soit diminuaient l'excentricité de l'ellipse présentée à l'écran avec les pas suivants : $\{-0.1; -0.01; -0.005\}$ soit les autres l'augmentaient avec les pas suivants : $\{0.01; 0.02; 0.15\}$. De plus, leurs rôles, (i.e. augmenter ou diminuer l'excentricité), étaient également contrebalancés à travers les sujets. Pour chaque taille, 11 répétitions ont été effectuées avec 11 valeurs initiales pour chaque taille, ces valeurs étant également réparties entre 0 et 1. Le test était donc composé de 33 essais, i.e. 3 (tailles) \times {11 répétitions}, présentés selon deux séries pseudo-aléatoires contrebalancées entre les sujets. Les sujets étaient incités à parcourir l'ensemble des possibilités à l'aide des flèches *les plus grandes* puis à encadrer leur choix avec les flèches *les plus petites*.

Analyse des données

Pour chaque taille et chaque sujet, 11 excentricités finales étaient collectées puis moyennées. Une analyse intra-sujet a été réalisée afin d'évaluer la variabilité du prototype (s'il en est un). Ainsi, une régression logistique avec un modèle LOGIT a été effectuée pour les résultats de chaque sujet afin de déterminer la médiane et le seuil de discrimination (JND³) pour chaque taille. Cette analyse avait pour objectif d'évaluer si chaque sujet a bien calibré les ellipses autour d'un prototype et dans quelle mesure ce prototype est commun à travers les sujets. En pratique, pour chaque sujet et chaque taille, une fonction de répartition (densité de probabilité cumulée) a été établie à partir des histogrammes des valeurs finales des excentricités. Les fonctions de répartitions ajustées présentent une forme de sigmoïde et ont été ajustées avec le modèle suivant, qui est un cas particulier du modèle LOGIT (voir équation 3.7).

En effet, il s'est avéré que les excentricités moyennes finales étaient centrées autour d'une valeur proche d'une borne de l'intervalle d'excentricités ajustées (vers .9, voir dans la partie résultats). Afin d'ajuster de façon satisfaisante la fonction de répartition (i.e. avec un point d'inflexion à peu près au milieu de l'intervalle de données), les données qui ont été ajustées sont les phases relatives ϕ du système 3.1 que l'on peut extraire de l'excentricité à partir de la formule 3.6. En effet, la phase relative ϕ correspondant à une excentricité de .9 est environ de 45 degrés ce qui correspond donc au milieu de l'intervalle de phases relatives possibles (i.e. entre 0 et 90 degrés). Cette transformation permet donc d'ajuster le modèle avec une bonne précision aux bornes.

$$LOGIT(\phi) = \frac{1}{1 + \exp\{-(\phi - \phi_{med})S\}} \quad (3.7)$$

où ϕ_{med} correspond à la valeur médiane des observations et S à la pente de la sigmoïde au point d'inflexion. Le seuil de discrimination (JND) peut ainsi être déterminé très simplement à partir du modèle ajusté. Il correspond à la différence entre la phase relative médiane et la phase relative dont

3. JND : Just Noticeable Difference.

la probabilité prédite par le modèle ajusté correspond à 75%. La figure 3.2 schématise les différentes grandeurs utilisées dans l'analyse. Pour chaque ajustement, la qualité de l'ajustement a été évaluée en testant la normalité des résidus, de plus la significativité de chaque coefficient a été testée avec la statistique de Fisher⁴.

Afin d'évaluer l'effet de la taille sur la valeur médiane et les JND de la phase relative calibrée, des ANOVA à mesures répétées ont été réalisées sur les médianes et JND avec pour unique facteur la taille. Le risque était défini à $\alpha = .05$, et les effets significatifs analysés plus en détail avec un test de Newman-Keuls.

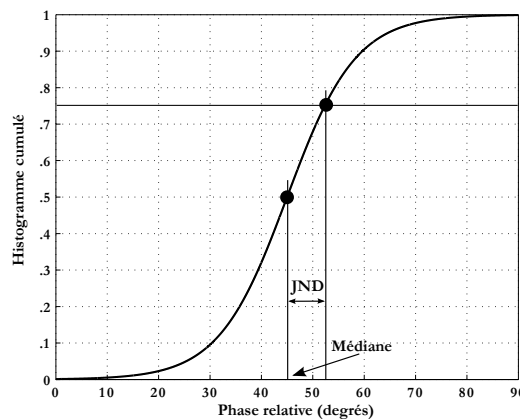


FIGURE 3.2 – Exemple de médiane et JND calculés de la régression logistique

3.2.2 Résultats

La figure 3.3 présente les fonctions de répartition ajustées par sujets et la fonction de répartition ajustée sur l'ensemble des résultats de toute la population. La table 3.2 présente les valeurs moyennes des médianes et JND par taille ainsi que les valeurs de la régression à l'échelle de la population globale. Les résultats détaillés sujet par sujet sont présentés dans la table 7.3 en annexe 7.3. Les coefficients des régressions sont significatifs pour chaque sujet et pour la population globale. De plus les corrélations entre les modèles ajustés et les données sont très élevées et toutes significatives ($r^2 > .9$, $p < .001$ pour tous les sujets).

Les ANOVA à mesures répétées sur les médianes et JND ne mettent en évidence aucun effet de la taille sur les médianes ($F(2, 38) = 0.687$, $p = 0.498$) et sur le JND ($F(2, 38) = 0.373$, $p = 0.665$). Ce résultat met donc en évidence l'existence d'un prototype commun à travers les sujets. Ce prototype a une valeur médiane moyenne proche de .91 (i.e. 45 degrés). L'analyse à l'échelle de la population met en évidence une valeur médiane et JND moyens du même ordre que ceux mis en évidence à l'échelle des sujets ($\phi_{med} \approx 45$ degrés et $JND \approx 7$ degrés).

3.2.3 Discussion intermédiaire

Dans cette expérience les sujets devaient ajuster l'excentricité d'une ellipse afin qu'elle leur évoque la forme la plus représentative de cette catégorie de forme géométrique. Différentes tailles d'ellipses ont été testées. L'existence d'un prototype visuel statique de l'ellipse dont la valeur de l'excentricité ne dépend pas de la taille et a une valeur médiane de .91 a été mise en évidence. Ce résultat s'ajoute aux résultats déjà connus sur l'existence de prototypes géométriques caractérisant une catégorie de forme géométrique (Feldman, 2000; Kalénine *et al.*, 2013; Theurel *et al.*, 2012). Les proportions de cette ellipse vont également dans le sens des résultats déjà mis en évidence par Wamain *et al.* (2011). En

4. L'ajustement et les tests statistiques pour la régression logistique ont été effectués en utilisant la fonction *NonLinearModel.fit* de Matlab.

Taille		Médiane			JND			Population	
		Moyenne	CI95% _{o_{inf}}	CI95% _{o_{sup}}	Moyenne	CI95% _{o_{inf}}	CI95% _{o_{sup}}	Médiane	JND
A1	ϕ	45.83	44.28	45.39	6.513	5.881	7.145	45.7	7.28
	e	.906	.898	.914	—	—	—	.906	—
A2	ϕ	44.17	42.92	43.23	6.655	6.110	7.200	45.5	7.52
	e	.908	.907	.918	—	—	—	.908	—
A3	ϕ	47.49	45.64	47.55	6.965	6.103	7.827	44.0	7.03
	e	.913	.907	.919	—	—	—	.915	—

TABLE 3.2 – Médianes et JND moyennées à travers les sujets et valeurs de la régression à travers l’ensemble des résultats de la population. Les intervalles de confiance à 95% sont également présentés pour les résultats moyennés à travers les sujets. Ils mettent clairement en évidence la consistance de la variabilité intra-sujet, aussi bien au niveau des médianes que des JND.

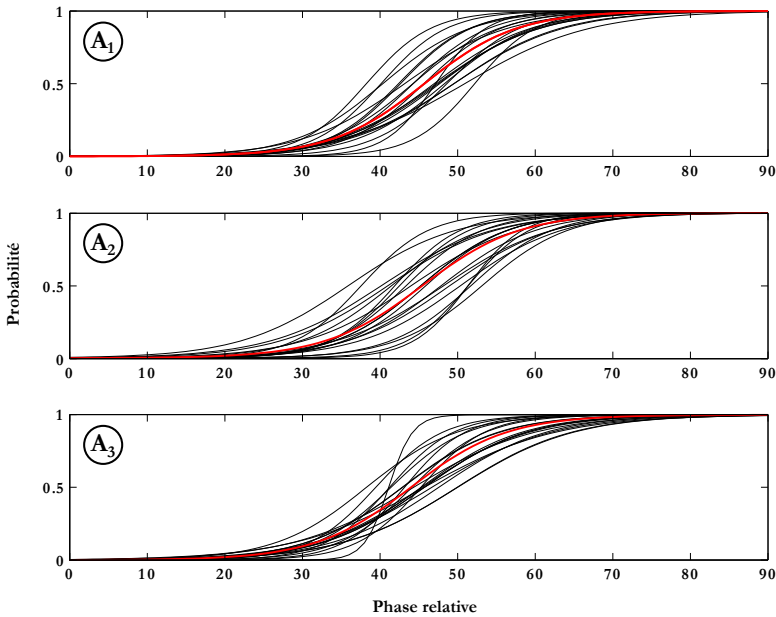


FIGURE 3.3 – Fonctions de répartitions ajustées pour chaque sujet (courbes noires) et à travers l’ensemble des résultats (courbe rouge) pour l’expérience 4

effet, dans cette étude les auteurs ont montré que nous discriminons visuellement plus facilement les ellipses dont les proportions correspondaient à celles que l'on dessine naturellement. Bien que cette étude ait soulevée une variabilité réduite dans la géométrie des productions motrices, ces ellipses avaient toutes une excentricité proche de .9. Le résultat de notre expérience semble donc soutenir l'hypothèse que le prototype visuel correspond également à une dynamique motrice préférentielle. Cependant, dans leur étude, Wamain *et al.* (2011) n'ont pas explicitement demandé aux sujets de dessiner l'ellipse qui leur paraissaient être la plus prototypique. La géométrie des ellipses dessinées a donc pu être uniquement contrainte par les caractéristiques motrices, et on a vu en introduction de ce chapitre et dans l'état de l'art que la valeur .9 semblait être un attracteur moteur des mouvements elliptiques (Athenes *et al.*, 2004). La forme la plus prototypique ne correspond pas nécessairement à celle que l'on dessine la plus naturellement, des facteurs esthétiques pouvant également entrer en jeu. La prochaine expérience a donc pour objectif d'évaluer s'il existe également une ellipse prototypique du point de vue de la production graphique et si cette ellipse est dans les mêmes proportions que celles mises en évidence du point de vue visuel.

3.3 Expérience 5 - Prototype moteur

Dans cette expérience, le but est d'évaluer dans une tâche de production motrice s'il existe également un prototype moteur de l'ellipse, et si tel est le cas, s'il a la même excentricité que celui mis en évidence dans le cas visuel statique.

3.3.1 Méthodes

Participants

20 sujets (5 femmes) d'un âge moyen de 39.7 ans ($SD = 11.52$) ont pris part volontairement à ce test. Tous les sujets étaient naïfs au sujet de l'expérience et avaient une vision normale ou corrigée. Tous les sujets ont donné leur consentement écrit avant de prendre part au test. Aucun des sujets n'a participé à l'expérience 4.

Appareillage

Les sujets étaient assis devant un écran d'ordinateur DELL 1907fp dont la résolution était de 1280×1024 pixels et devant une tablette graphique Wacom Intuos5 permettant d'enregistrer les données à une fréquence de 129 Hz et avec une résolution spatiale de 5.10^{-3} mm. L'interface d'enregistrement a été programmée avec le logiciel Max/MSP.

Tâche

Dans cette expérience, il était demandé aux sujets de dessiner en continu une ellipse sur une tablette graphique pendant 50 secondes. Les sujets pouvaient voir leurs mains pendant qu'ils dessinaient. L'expérience était composée de trois sessions de 50 secondes au cours desquelles les sujets devaient dessiner soit une petite ellipse, soit une ellipse de taille moyenne, soit une grande ellipse. L'ordre dans lequel les sujets dessinaient ces trois types d'ellipse a été contrebalancé entre les sujets. Aucun patron des différentes tailles n'était fourni aux sujets et ils avaient la possibilité de s'entraîner avant le début du test afin de choisir les trois tailles d'ellipses. Il leur était explicitement précisé comme dans les autres expériences que l'ellipse qu'ils allaient dessiner devait être la plus représentative de cette catégorie de forme géométrique. De plus, afin de minimiser la variabilité de leurs trajectoires, il était demandé aux sujets d'être le plus précis possible spatialement au cours de leurs mouvements.

Analyse des données

Pour chaque sujet, trois enregistrements ont été réalisés (un par taille). Chacun correspond aux coordonnées du stylet sur la tablette graphique échantillonnées au cours du temps (x_n, y_n). Les données brutes ont été lissées afin d'éliminer le bruit numérique lié au taux d'échantillonnage élevé en appliquant un filtre de Savitsky-Golay (Savitzky et Golay, 1964) avec une fenêtre de 43 échantillons et une interpolation du troisième ordre, ce qui est équivalent à un filtrage passe-bas avec une fréquence de coupure de 8 Hz. De plus, un filtrage passe-haut (Butterworth) avec une fréquence de coupure de 0.2 Hz a été appliqué afin d'éliminer la dérive spatiale des sujets au cours du mouvement. Afin de déterminer les caractéristiques géométriques des mouvements, et en particulier l'excentricité de l'ellipse dessinée, nous avons utilisé le modèle de geste présenté précédemment en considérant que le mouvement était le résultat de l'oscillation de deux oscillateurs harmoniques déphasés d'une phase relative ϕ . La figure 3.4 présente un exemple d'enregistrement et la phase relative associée.

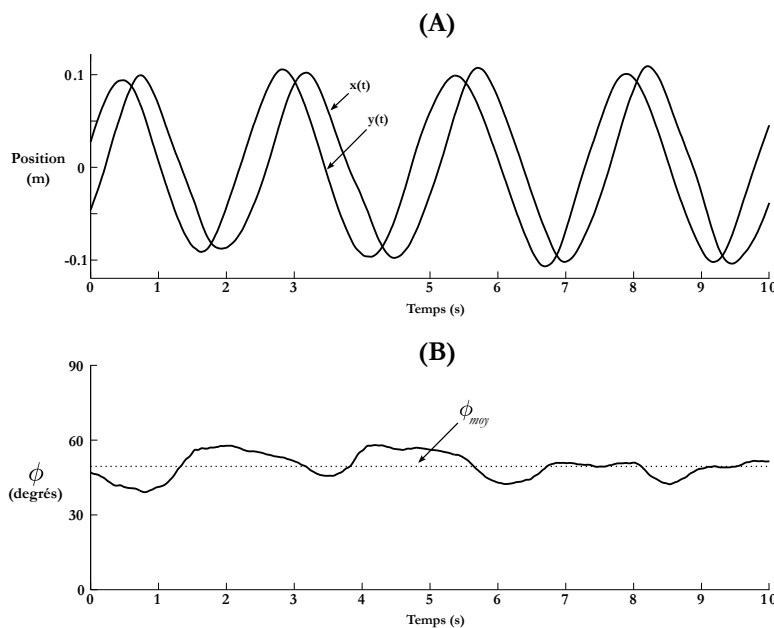


FIGURE 3.4 – Exemple d'enregistrement et de phase relative associée pour une production motrice typique d'un sujet. Le panel A présente les coordonnées au cours du temps et on voit clairement le déphasage quasiment constant entre $x(t)$ et $y(t)$, le panel B présente la phase relative associée au cours temps.

Pour chaque enregistrement, deux séries $x(t)$ et $y(t)$ ont été collectées et l'objectif était donc d'estimer la valeur de ϕ au cours du temps. Une méthode facile pour estimer le déphasage au cours du temps entre deux oscillateurs est d'utiliser la transformée de Hilbert (Panter, 1965; Smith et Merseureau, 1991). En pratique pour calculer $\phi(t)$, on calcule la différence entre les phases déroulées des transformées de Hilbert \tilde{x} et \tilde{y} de x et y , soit $\phi(t) = |\tilde{x}(t) - \tilde{y}(t)|$. Une fois la phase relative calculée, elle a été moyennée (ϕ_{moy}) pour les analyses. Enfin, l'excentricité a été déterminée à partir de la formule $e_{moy} = \sqrt{1 - \tan^2\left(\frac{\phi_{moy}}{2}\right)}$.

Comme pour l'expérience 4, des régressions logistiques sont réalisées pour chaque sujet et pour la population globale dans le but d'évaluer l'existence et la variabilité des productions et du prototype (s'il en est un) pour chacun des sujets. Des ANOVA à mesures répétées sont réalisées sur ces deux descripteurs avec pour unique facteur la taille (A_1 , A_2 et A_3). La normalité des échantillons a été systématiquement vérifiée avec un test de Lilliefors.

3.3.2 Résultats

La figure 3.5 présente les fonctions de répartitions ajustées par sujet et la fonction de répartition ajustée sur l'ensemble des résultats de toute la population pour chaque taille. La table 3.3 présente les valeurs moyennes des médianes et JND par taille, et également les valeurs de la régression à l'échelle de la population globale. Les résultats détaillés sujet par sujet sont présentés dans la table 7.4 en annexe 7.3. Les coefficients des régressions sont significatifs pour chaque sujet et pour la population globale, de plus les corrélations entre les modèles ajustés et les données sont très élevées et toutes significatives ($r^2 > .9$, $p < .001$ pour tous les sujets).

Les ANOVA à mesures répétées sur les médianes ne mettent en évidence aucun effet de la taille ($F(2, 38) = 1.3$, $p = 0.285$). Ce résultat met donc en évidence l'existence d'un prototype commun à travers les sujets avec une excentricité proche de .91 (i.e. 45 degrés). Le JND est par contre significativement modifié en fonction de la taille ($F(2, 38) = 9.13$, $p < .001$), en particulier, les JND ont été plus importants pour la *grande* ellipse que pour la petite et la moyenne (Newman-Keuls : $p < .01$ pour la petite et $p < .001$ pour la moyenne). Ce prototype a une valeur médiane moyenne proche de .91 (i.e. 45 degrés). Cela révèle que la variabilité intra-sujet a été plus grande pour cette ellipse qui a donc été dessinée de façon moins précise. L'analyse à l'échelle de la population met en évidence que la médiane est bien cohérente avec la moyenne des médianes intra-sujet, et qu'à l'inverse, la variabilité à l'échelle de la population est à peu près doublée par rapport à la variabilité intra-sujet (9 degrés à l'échelle de la population contre 4 degrés pour la variabilité intra-sujet). Cela est clairement visible sur la figure 3.5.

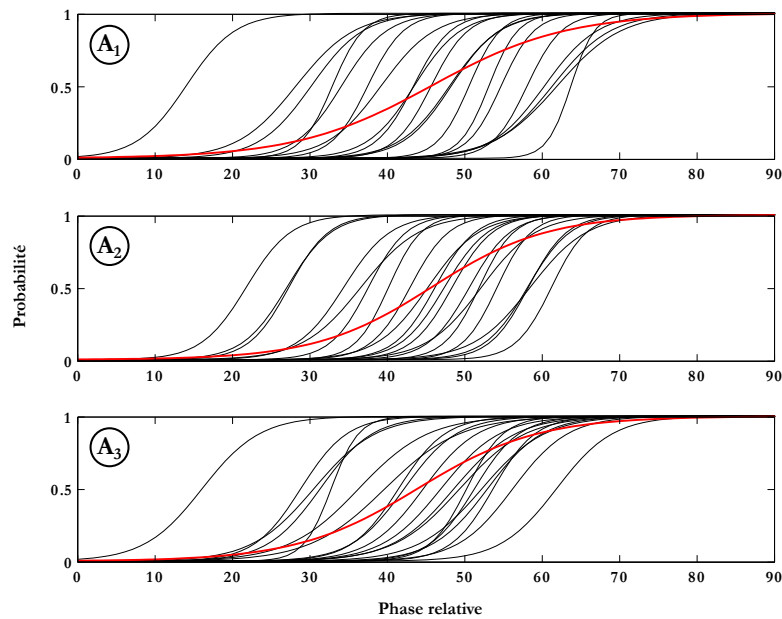


FIGURE 3.5 – Fonctions de répartitions ajustées pour chaque sujet (courbes noires) et à travers l'ensemble des résultats (courbe rouge) pour l'expérience 5

3.3.3 Discussion intermédiaire

Dans cette expérience, les sujets devaient dessiner trois ellipses continûment pendant 50 secondes, une petite, une moyenne, et une grande, et ces ellipses devaient être les plus représentatives de la catégorie de forme géométrique ellipse. Les résultats ont mis en évidence l'existence d'un prototype caractérisé par l'excentricité de l'ellipse dessinée indépendamment de la taille de l'ellipse. La médiane moyennée à travers les sujets est de .91 et le prototype semble donc cohérent avec celui mis en évidence dans l'expérience 4 (une comparaison détaillée sera effectuée après la présentation des

Taille		Médiane			JND			Population	
		Moyenne	CI95% _{inf}	CI95% _{sup}	Moyenne	CI95% _{inf}	CI95% _{sup}	Médiane	JND
A1	ϕ	44.92	43.28	45.37	3.5295	3.167	3.891	45.6	10.36
	e	.908	.877	.933	—	—	—	.907	—
A2	ϕ	39.56	38.33	39.92	3.5298	3.298	3.761	45.6	8.3
	e	.910	.894	.931	—	—	—	.907	—
A3	ϕ	49.91	48.22	51.18	4.2683	3.900	4.636	43.8	9.17
	e	.917	.894	.937	—	—	—	.916	—

TABLE 3.3 – Médianes et JND moyennées à travers les sujets et valeurs de la régression à travers l'ensemble des résultats de la population. Les intervalles de confiance à 95% sont également présentés pour les résultats moyennés à travers les sujets, ils mettent clairement en évidence la consistance de la variabilité intra-sujet, aussi bien au niveau des médianes que des JND.

quatre expériences). Ce résultat va dans le sens des observations réalisées par Athenes *et al.* (2004); Sallagoity *et al.* (2004); Danna *et al.* (2011) sur l'existence d'une ellipse d'excentricité préférentielle. Dans cette expérience, il était explicitement demandé aux sujets de dessiner l'ellipse qui leur paraissait la plus prototypique, ce qui n'était pas le cas dans les expériences de Wamain *et al.* (2011) qui ne s'intéressaient qu'à la production motrice sans contrôle de l'intention des sujets selon un critère particulier, ici la prototypicité de la forme.

Il est cependant nécessaire de pondérer notre conclusion au regard de la variabilité intra-sujet, qui est en effet plus élevée que dans l'expérience 4, cela ne remet pas en cause le fait qu'à l'échelle de la population un prototype semble émerger dont la valeur est proche de .9. Cependant, on peut faire l'hypothèse que la variabilité observée est plus grande en raison de caractéristiques biomécaniques différentes à travers les sujets, cette variabilité avait également été observée dans les expériences de Wamain *et al.* (2011), qui avaient mis en évidence le caractère idiosyncrasique de l'ellipse préférentielle.

Malgré cette pondération, on peut néanmoins affirmer qu'une ellipse d'excentricité proche de .9 semble bien caractériser cette catégorie de forme géométrique au niveau moteur. Les deux premières expériences ont donc montré l'existence d'un prototype géométrique pour l'ellipse aussi bien du point de vue visuel statique que du point de vue moteur. Cependant, notre hypothèse initiale est que le prototype visuel émerge de la dynamique du geste, bien que les deux expériences précédentes semblent le confirmer, mais on ne peut pas encore trancher sur le fait que ce sont les caractéristiques motrices qui ont impacté les caractéristiques visuelles du prototype visuel ou si c'est l'inverse, i.e. que les caractéristiques motrices se soient adaptées aux caractéristiques des ellipses que l'on rencontre le plus dans notre environnement.

3.4 Expérience 6 - Prototype visuel dynamique

Afin d'évaluer le caractère dynamique du prototype de l'ellipse mis en évidence précédemment, et plus spécifiquement si le prototype émerge d'une dynamique motrice, l'expérience 6 vise donc à voir s'il existe un prototype visuel dynamique de l'ellipse et si il a des caractéristiques similaires aux deux situations déjà testées. Ainsi, l'expérience 4 va être reproduite mais cette fois-ci avec des stimuli dynamiques (points lumineux en mouvement) générés à partir du modèle de geste précédent dont

TABLE 3.4 – Amplitudes en *cm* et périodes en secondes des ellipses dynamiques utilisées dans le test

A_1	A_2	A_3
3 cm	5 cm	10 cm
T_1	T_2	T_3
1.2 s	1.5 s	1.8 s


la dynamique correspond à celle d'un mouvement humain, i.e. respectant la loi en puissance 1/3.

3.4.1 Méthodes

Participants

20 sujets (3 femmes) d'un âge moyen de 32 ans ($SD = 11.42$) ont pris part volontairement à ce test. Tous les sujets étaient naïfs au sujet de l'expérience et tous avaient une vision normale ou corrigée. Tous ont donné leur consentement écrit avant de prendre part au test. Aucun des sujets n'a participé aux expériences 4 & 5.

Stimuli

 Les stimuli visuels utilisés dans cette expérience sont composés de points lumineux parcourant dynamiquement une ellipse à différentes fréquences (point blanc sur fond noir). Différentes tailles et différentes excentricités ont été considérées et générées à partir du modèle de geste présenté précédemment. Ces ellipses sont donc définies par :

$$\begin{cases} x(t) = A \cos(\omega t) \\ y(t) = A \cos(\omega t + \phi) \end{cases} \quad (3.8)$$

avec A l'amplitude, $\omega = \frac{2\pi}{T}$ où T est la période du mouvement. L'excentricité de l'ellipse e peut donc être aisément contrôlée en utilisant la relation $\phi = 2 \arctan \sqrt{1 - e^2}$ mise en évidence précédemment. Ainsi, l'excentricité de l'ellipse peut être modifiée en manipulant implicitement la phase relative ϕ . Comme pour l'expérience 4, il est important de noter que l'inclinaison de l'ellipse était toujours orientée de 45 degrés par rapport à l'horizontale. Les tailles choisies sont les mêmes que dans l'expérience 4. Trois ellipses de tailles différentes (A_1 , A_2 et A_3) ont été utilisées dans ce test. De plus, trois fréquences de geste (de périodes T_1 , T_2 et T_3) ont été utilisées. La table 3.4 résume les tailles et périodes utilisées.

Appareillage

Les sujets étaient assis devant un écran DELL 1907fp de résolution 1280 x 1024 et de fréquence de rafraichissement 60 Hz. Les ellipses dynamiques étaient affichées au centre de l'écran et les sujets pouvaient ajuster leurs excentricités par le biais d'un clavier MIDI AKAI MPK. L'interface de test a été programmée avec le logiciel Max/MSP.

Tâche

La tâche et le protocole d'ajustement utilisés dans cette expérience étaient les mêmes que dans l'expérience 4. Il était demandé aux sujets d'ajuster l'excentricité de l'ellipse dynamique afin de régler celle qui leur paraissait la plus représentative de cette catégorie de formes géométriques. Pour chaque combinaison (taille) x (fréquence), 11 répétitions ont été effectuées avec 11 valeurs initiales pour chaque taille, ces valeurs étant également réparties entre 0 et 1. Le test était donc composé de

99 essais, i.e. 3 (tailles) \times 3 (fréquences) \times {11 valeurs initiales}, présentés selon deux séries pseudo-aléatoires contrebalancées entre les sujets. Les sujets étaient incités à parcourir l'ensemble des possibilités à l'aide des flèches *les plus grandes* puis à encadrer leurs choix avec les flèches *les plus petites*.

Analyse des données

Pour chaque couple (taille) \times (fréquence) et chaque sujet, 11 excentricités finales ont été collectées puis transformée en phase relative à partir de la formule 3.3. Comme pour les deux expériences précédentes, des régressions logistiques ont été réalisées pour chaque sujet et pour la population globale dans le but d'évaluer l'existence et la variabilité des productions et du prototype (s'il en est un) pour chacun des sujets et pour chaque couple (taille) \times (fréquence). Des ANOVA à mesures répétées sur les médianes et JND ont été réalisées avec pour facteur la taille et la fréquence. Dans tous les tests, le risque était défini à $\alpha = .05$ et les effets significatifs analysés avec des tests post-hoc de Newman-Keuls. De même que dans les deux expériences précédentes, une analyse intra-sujet a été réalisée. La normalité des échantillons a été préalablement vérifiée avec des tests de Lilliefors.

3.4.2 Résultats

La figure 3.6 présente les fonctions de répartitions ajustées par sujet et la fonction de répartition ajustée sur l'ensemble des résultats de toute la population pour chaque couple (taille) \times (fréquence). La table 3.5 présente les valeurs moyennes des médianes et JND par taille et fréquence et également les valeurs de la régression à l'échelle de la population globale. Les résultats détaillés sujet par sujet sont présentés dans la table 7.5 en annexe 7.3. Les coefficients des régressions sont significatifs pour chaque sujet et pour la population globale. De plus les corrélations entre les modèles ajustés et les données sont très élevées et toutes significatives ($r^2 > .9$, $p < .001$ pour tous les sujets).

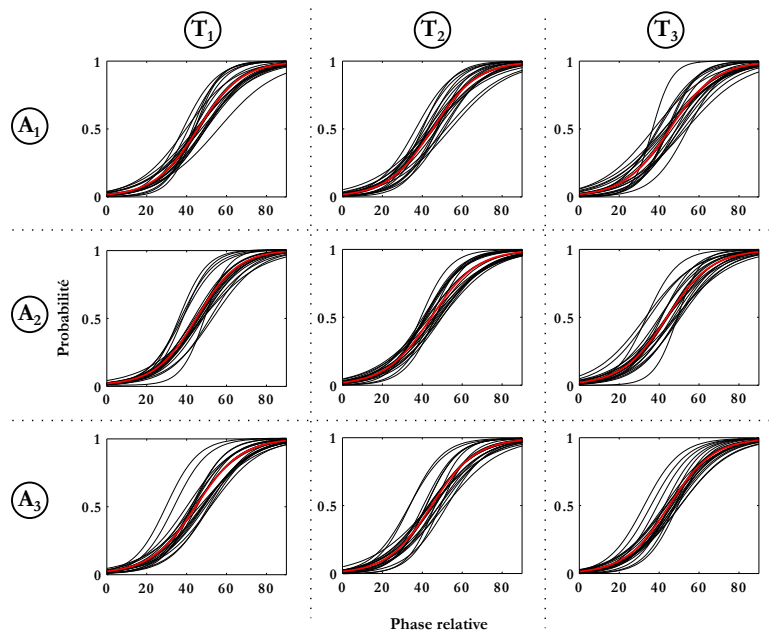


FIGURE 3.6 – Fonctions de répartitions ajustées pour chaque sujet (courbes noires) et à travers l'ensemble des résultats (courbe rouge) pour l'expérience 6 par taille et par fréquence

Les ANOVA à mesures répétées sur les médianes et JND ne mettent en évidence aucun effet de la taille, ni de la fréquence, ni même d'interaction sur les médianes (Taille : $F(2, 38) = 0.951$, $p = 0.392$, Fréquence : $F(2, 38) = 0.854$, $p = 0.432$, (Taille) \times (Fréquence) : $F(4, 76) = 0.171$, $p = 0.919$) et sur le JND (Taille : $F(2, 38) = 0.603$, $p = 0.552$, Fréquence : $F(2, 38) = 0.682$, $p = 0.51$, (Taille) \times (Fréquence) : $F(4, 76) = 0.682$, $p = 0.51$). Ce résultat met donc en évidence l'existence d'un prototype

commun à travers les sujets. Ce prototype a une valeur médiane moyenne proche de .91 (i.e. 45 degrés). L'analyse à l'échelle de la population met en évidence une valeur médiane et JND moyens du même ordre que ceux mis en évidence à l'échelle des sujets ($\phi_{med} \approx 45$ degrés et JND ≈ 12 degrés).

Taille	Fréquence		Médiane			JND			Population	
			Moyenne	CI95% _{inf}	CI95% _{sup}	Moyenne	CI95% _{inf}	CI95% _{sup}	Médiane	JND
A1	T1	ϕ	45.78	44.03	47.54	12.410	11.281	13.538	45.6	12.41
		e	.906	.914	.914	—	—	—	.907	—
	T2	ϕ	45.06	43.12	46.99	12.589	11.454	13.725	44.8	13.18
		e	.909	.900	.918	—	—	—	.911	—
	T3	ϕ	45.53	43.28	47.77	12.768	11.399	14.136	45.3	12.6
		e	.908	.896	.917	—	—	—	.910	—
A2	T1	ϕ	45.32	43.26	47.37	11.528	10.490	12.565	45.1	11.92
		e	.908	.898	.933	—	—	—	.901	—
	T2	ϕ	44.42	42.96	45.88	12.877	11.797	13.958	44.2	12.82
		e	.912	.906	.919	—	—	—	.914	—
	T3	ϕ	43.79	41.47	46.10	12.508	11.529	13.487	43.8	13.20
		e	.915	.904	.925	—	—	—	.916	—
A3	T1	ϕ	44.86	42.38	47.33	12.438	11.545	13.330	44.68	13.31
		e	.910	.898	.921	—	—	—	.912	—
	T2	ϕ	44.34	42.15	46.53	12.303	11.194	13.412	44.19	12.81
		e	.913	.902	.922	—	—	—	.914	—
	T3	ϕ	44.86	42.80	46.91	12.088	11.213	12.963	44.8	12.18
		e	.910	.900	.920	—	—	—	.911	—

TABLE 3.5: Résultats de l'expérience 6 – Moyennes des médianes et JND

3.4.3 Discussion intermédiaire

Dans cette expérience les sujets devaient calibrer l'excentricité d'une ellipse dynamique (i.e. un point lumineux parcourant une trajectoire elliptique) dont la taille et la fréquence du mouvement variaient. Le mouvement était généré à partir d'un modèle de geste qui considère qu'un mouvement graphique elliptique est le résultat du couplage entre deux oscillateurs harmoniques de même fréquence et de même amplitude, déphasés par une phase relative ϕ . Les résultats ont mis en évidence l'existence d'un prototype caractérisé par l'excentricité de l'ellipse dessinée indépendamment de la taille de l'ellipse et de la fréquence du mouvement. La valeur médiane moyennée à travers les sujets est de .91, ce prototype est donc cohérent avec celui mis en évidence dans les expériences 4 et 5 (une comparaison détaillée sera effectuée après la présentation des quatre expériences).

Finalement, ces trois premières expériences ont permis de montrer que les ellipses sont caractérisées par un prototype visuel statique d'excentricité proche .9, que ce prototype est cohérent avec le prototype moteur originel, et que ce prototype visuel est cohérent quand le stimulus visuel est dynamique et respecte une dynamique motrice. On peut donc conclure que la géométrie du prototype est cohérente avec une dynamique motrice, cependant dans la dernière expérience les sujets se sont probablement imaginés une forme géométrique à partir du point lumineux, et ont ainsi pu se baser à la fois sur la géométrie de la trajectoire parcourue par le point et par la dynamique. On ne peut donc pas être certain de l'existence d'un prototype dynamique, et nous ne pouvons pas décider si la géométrie du prototype est une conséquence d'une dynamique caractéristique des mouvements biologiques caractérisée elle par la phase relative entre deux oscillateurs harmoniques.

3.5 Expérience 7 - Prototype sonore


Afin de déterminer si le prototype géométrique de l'ellipse émerge d'une expérience visuelle ou si les caractéristiques géométriques de ce prototype sont une conséquence de contraintes biomécaniques dynamiques, ici d'attracteurs moteurs, il est nécessaire de pouvoir évaluer si la dynamique correspondant à une ellipse d'excentricité .9 est prototypique en soi. La modalité visuelle ne permet pas d'évoquer la dynamique du mouvement indépendamment d'une trajectoire, et donc d'une forme géométrique spécifique. Une modalité qui semble bien adaptée pour ce genre de tâche est la modalité auditive. En effet, on a vu dans les chapitres précédents qu'elle permettait d'évoquer des mouvements, et en particulier des mouvements graphiques sans pour autant utiliser d'indices spatiaux. En utilisant les variations de timbre générées par le son de frottement produit par une personne qui dessine, on peut ainsi générer des stimuli qui n'évoquent que la dynamique d'un mouvement sans le moindre indice géométrique (i.e. les sujets ne voyaient aucune forme, ils ne pouvaient utiliser que le geste évoqué par le son pour s'imaginer une géométrie). La prochaine expérience a donc pour objectif d'évaluer, par le biais de la modalité auditive, si la dynamique du geste est prototypique en soi indépendamment de la géométrie de la forme. Pour cela, l'expérience précédente va être reproduite en *remplaçant* le point lumineux par un son de frottement de synthèse correspondant au son produit par le mouvement défini précédemment.

3.5.1 Méthodes

Participants

20 sujets (9 femmes) d'un âge moyen de 31 ans ($SD = 10.21$) ont pris part volontairement à ce test. Tous les sujets étaient naïfs au sujet de l'expérience, tous avaient une audition normale et ont tous donné leur consentement écrit avant de prendre part au test. Aucun des sujets n'a participé aux expériences 4, 5, & 6.

Stimuli

 Les stimuli sonores utilisés dans cette expérience sont composés de sons de frottement de synthèse générés à partir d'un profil de vitesse calculé avec le même modèle de geste que pour les expériences 4 & 6. En pratique le son est synthétisé en filtrant un bruit blanc passe-bande (Butterworth d'ordre 2 ; $Q = 5$) et en contrôlant la fréquence centrale du filtre par la vitesse du mouvement. Le mapping entre la vitesse et la fréquence est choisi arbitrairement à 20, i.e. $f_c = 20v_t$, où v_t est la vitesse tangentielle définie par l'équation 3.4. Les fréquences et tailles utilisées sont les mêmes que dans les expériences 4 & 6. On pourra noter que les différentes amplitudes ont finalement le même

effet qu'un changement de mapping. Ces ellipses sont donc définies par :

$$\begin{cases} x(t) = A \cos(\omega t) \\ y(t) = A \cos(\omega t + \phi) \end{cases} \quad (3.9)$$

avec $\omega = \frac{2\pi}{T}$ où T est la période du mouvement. L'excentricité de l'ellipse e peut donc être aisément contrôlée en utilisant la relation $\phi = 2 \arctan \sqrt{1 - e^2}$ mise en évidence précédemment. Ainsi, l'excentricité de l'ellipse peut être modifiée en manipulant implicitement la phase relative ϕ . Les tailles choisies sont les mêmes que dans l'expérience 4, trois ellipses de tailles différentes (i.e. l'amplitude A utilisée dans le système 3.5) ont été utilisées dans ce test. De plus, trois fréquences de geste ont été utilisées. La table 3.4 résume les tailles et périodes utilisées.

Appareillage

Les sujets étaient assis dans une cabine audiométrique, les sons étaient diffusés par le biais d'un casque Sennheiser HD-650 à une fréquence d'échantillonnage de 44100 Hz. Les sujets pouvaient ajuster le son à l'aide d'un clavier MIDI AKAI MPK. Le volume a été ajusté pour procurer un niveau d'écoute suffisant et confortable.

Tâche

La tâche et le protocole d'ajustement utilisés dans cette expérience étaient les mêmes que dans les expériences 4 & 6. Dans cette expérience, il était demandé aux sujets de calibrer le son de frottement qui évoque une personne qui dessine une forme elliptique qui leurs parait la plus représentative de cette catégorie de formes géométriques. On insiste sur le fait qu'aucune forme visuelle n'était présentée aux sujets. Ils modifiaient implicitement l'excentricité de l'ellipse, ce qui modifiait le son de frottement qui était généré à partir du profil de vitesse du mouvement correspondant à l'équation 3.4. Cette tâche étant *a priori* moins intuitive que les tâches visuelles, il était précisé au sujet qu'il devait s'imaginer le geste à partir du son, puis régler ce geste afin qu'il leur évoque le mouvement le plus elliptique. Pour chaque couple (taille) x (fréquence), 11 répétitions ont été effectuées avec 11 valeurs initiales pour chaque taille. Ces valeurs étant également réparties entre 0 et 1. Le test était donc composé de 99 essais, i.e. 3 (tailles) x 3 (fréquences) x {11 valeurs initiales}, présentés selon deux séries pseudo-aléatoires contrebalancées entre les sujets. Les sujets étaient incités à parcourir l'ensemble des possibilités à l'aide des flèches *les plus grandes* puis à encadrer leur choix avec les flèches *les plus petites*.

Analyse des données

L'analyse des données et les tests statistiques réalisés sont les mêmes que pour l'expérience 6.

3.5.2 Résultats

La figure 3.7 présente les fonctions de répartitions ajustées par sujet et la fonction de répartition ajustée sur l'ensemble des résultats de toute la population pour chaque couple (taille) x (fréquence). La table 3.6 présente les valeurs moyennes des médianes et JND par taille et par fréquence et également les valeurs de la régression à l'échelle de la population globale. Les résultats détaillés sujet par sujet sont présentés dans la table 7.6 en annexe 7.3. Les coefficients des régressions sont significatifs pour chaque sujet et pour la population globale, de plus les corrélations entre les modèles ajustés et les données sont très élevées et toutes significatives ($r^2 > .9$, $p < .001$ pour tous les sujets).

Les ANOVA à mesures répétées sur les médianes ainsi que les JND ne mettent en évidence aucun effet de la taille, ni de la fréquence, ni même d'interaction sur les médianes (Taille : $F(2, 38) =$

0.735, $p = 0.486$, Fréquence : $F(2, 38) = 0.22$, $p = 0.797$, (Taille) x (Fréquence) : $F(4, 76) = 0.579$, $p = 0.661$) et sur le JND (Taille : $F(2, 38) = 0.312$, $p = 0.703$, Fréquence : $F(2, 38) = 5.15$, $p = 0.0119$, (Taille) x (Fréquence) : $F(4, 76) = 0.117$, $p = 0.965$). Ce résultat met donc en évidence l'existence d'un prototype commun à travers les sujets. Ce prototype a une valeur médiane moyenne proche de .83 (i.e. $\phi = 58.3$ degrés). L'analyse à l'échelle de la population met en évidence une valeur médiane et JND moyens du même ordre que ceux mis en évidence à l'échelle des sujets ($\phi_{med} = 58$ degrés et JND = 11 degrés).

Taille	Fréquence		Médiane			JND			Population	
			Moyenne	CI95% _{inf}	CI95% _{sup}	Moyenne	CI95% _{inf}	CI95% _{sup}	Médiane	JND
A1	T1	ϕ	58.09	56.34	59.84	11.453	10.270	12.635	58.40	11.574
		e	.831	.817	.844	—	—	—	.829	
	T2	ϕ	58.70	56.94	60.45	10.249	9.046	11.452	58.60	10.370
		e	.826	.812	.840	—	—	—	.828	
	T3	ϕ	59.67	57.92	61.42	9.977	8.879	11.076	59.60	10.400
		e	.819	.804	.832	—	—	—	.820	
A2	T1	ϕ	58.45	55.53	61.36	11.350	10.123	12.576	59.00	11.970
		e	.828	.804	.850	—	—	—	.825	
	T2	ϕ	58.47	56.25	60.69	10.274	9.209	11.339	58.50	10.540
		e	.828	.810	.845	—	—	—	.828	
	T3	ϕ	58.11	56.46	59.76	10.133	9.106	11.159	58.30	9.700
		e	.831	.818	.843	—	—	—	.830	
A3	T1	ϕ	57.78	55.39	60.16	11.419	10.015	12.824	58.20	11.780
		e	.834	.815	.851	—	—	—	.831	
	T2	ϕ	58.24	56.43	60.04	10.909	9.883	11.935	58.30	11.710
		e	.830	.816	.843	—	—	—	.830	
	T3	ϕ	57.01	54.87	59.15	10.183	9.427	10.939	57.00	10.990
		e	.839	.823	.854	—	—	—	.834	

TABLE 3.6: Résultats de l'expérience 7 – Moyennes des médianes et JND

3.5.3 Discussion intermédiaire

Dans cette expérience, les sujets devaient calibrer le son de frottement produit par une personne qui dessine une ellipse afin que les variations de timbre leurs évoque la trajectoire elliptique la plus caractéristique de cette catégorie de forme géométrique. On a vu dans les expériences précédentes que l'ellipse d'excentricité .9 semble être caractéristique de cette catégorie de forme, aussi bien du point de vue visuel statique, moteur, que visuel dynamique. L'objectif de cette 4ème expérience était

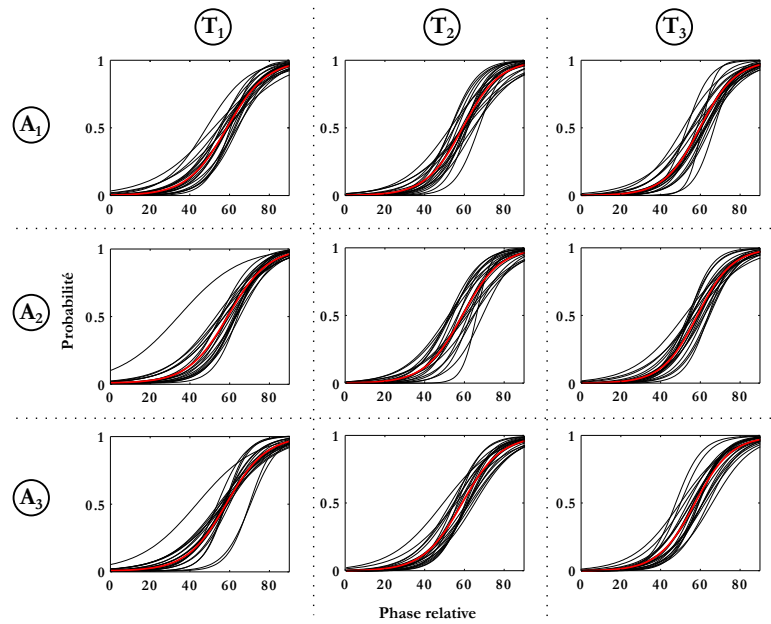


FIGURE 3.7 – Fonctions de répartitions ajustées pour chaque sujet (courbes noires) et à travers l’ensemble des résultats (courbe rouge) pour l’expérience 7 par taille et par fréquence

d’évaluer dans quelle mesure la dynamique seule du mouvement biologique correspondant à une ellipse d’excentricité .9 était prototypique en soi, ce qui suggérerait finalement que, la géométrie du prototype ne serait qu’une conséquence de contraintes dynamiques caractérisées par des attracteurs moteurs. Les sons ont été générés à partir d’un modèle de synthèse physiquement informé permettant de contrôler aisément la vitesse évoquée par les variations de timbre. Trois tailles d’ellipse et trois fréquences de geste ont été testées et sont les mêmes que celles de l’expérience 6.

Les résultats ont mis en évidence que la catégorie de forme ellipse est bien caractérisée du point de vue sonore par une excentricité particulière (proche de .8), et comme pour l’expérience visuelle dynamique, que cette excentricité ne dépend pas de la taille ou de la fréquence du geste. Cela montre donc qu’une dynamique est prototypique de cette catégorie de forme. Cependant, l’excentricité de l’ellipse prototypique associée ne semble pas correspondre à celle des expériences précédentes et a une médiane moyenne de .83. Le prototype sonore semble donc être *plus arrondi* que les prototypes mis en évidence dans les autres expériences et une comparaison détaillée va être réalisée dans la section suivante.

3.6 Comparaison entre les différents prototypes

Dans cette section, l’objectif est de comparer formellement les différents prototypes mis en évidence dans les quatre expériences précédentes. Dans la mesure où au sein de chaque expérience, aucun effet de la taille (expériences 4, 5, 6, & 7) ni fréquence (expériences 6 & 7) sur la valeur de l’excentricité du prototype n’a été mis en évidence, les différents prototypes seront comparés sans tenir compte de ces facteurs mais en ne considérant que l’expérience comme facteur. Dans un premier temps, nous comparerons l’excentricité des différents prototypes en fonction de la modalité, dans un second temps, nous nous intéresserons à la variabilité des prototypes, par sujet et au sein de la population globale, en fonction de la modalité. Il semble en effet que cette variabilité évolue entre les expériences et qu’elle soit différente si l’on considère la population globale ou bien les sujets indépendamment.

3.6.1 Excentricité

La table 3.7 présente les valeurs moyennes et erreurs standards des phases relatives médianes des prototypes pour chaque expérience. La figure 3.8 présente les quatre prototypes. En pratique, une ANOVA a été réalisée sur l’ensemble des phases relatives calibrées (expériences 4, 6, & 7) ou ajustées (Expérience 5) avec pour facteur le type d’expérience. La normalité des distributions a été vérifiée avant de réaliser l’ANOVA avec des tests de Lilliefors. Les résultats mettent en évidence une influence significative du type d’expérience sur la valeur moyenne de l’excentricité ($F(3, 476) = 188.92, p < .001$, voir figure 3.9). Un test post-hoc HSD de Tukey a mis en évidence que le prototype sonore est significativement plus arrondi que les trois autres prototypes ($p < .001$ pour toutes les comparaisons). Ces trois autres prototypes ne présentant pas de différences significatives.

	Visuel Statique	Moteur	Visuel Dynamique	Sonore
ϕ_{med} (degrés)	45.17 ± 1.01 ($e_{med} = .91$)	44.52 ± 2.99 ($e_{med} = .91$)	$44.94 \pm .67$ ($e_{med} = .91$)	$58.29 \pm .68$ ($e_{med} = .83$)

TABLE 3.7 – Valeurs moyennes et erreurs standards des phases relatives médianes des prototypes pour chaque expérience

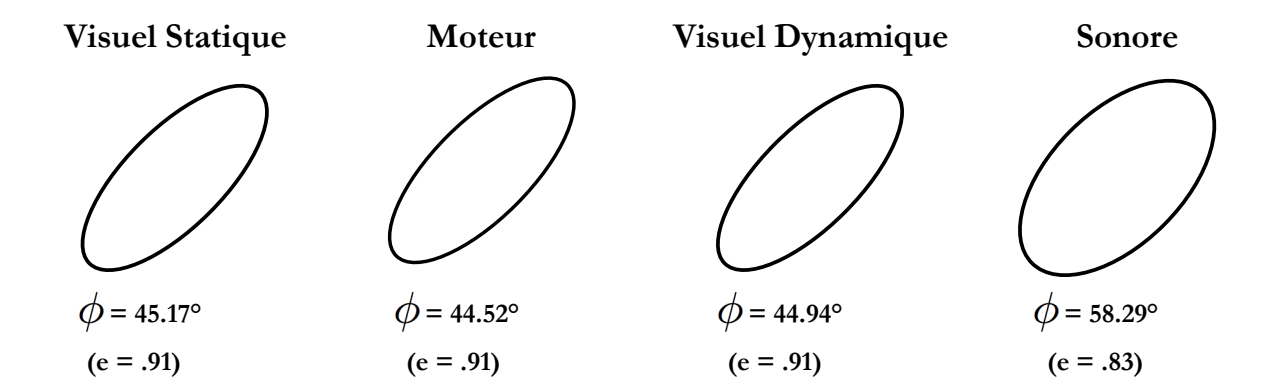


FIGURE 3.8 – Géométrie des prototypes mis en évidence dans les quatre expériences

3.6.2 Variabilité des prototypes entre les modalités

La table 3.8 présente les valeurs moyennes et écarts types des JND à travers et au sein des sujets pour les quatre expériences. Ces données permettent d’évaluer la variabilité des différents prototypes entre les modalités (sensorielles et sensori-motrice), mais également au sein des modalités (variabilité inter-sujet vs. variabilité intra-sujet).

Ainsi, on peut se rendre compte que le prototype Visuel Dynamique et le prototype Sonore ont des variabilités inter-sujet plus grandes que celles des prototypes Visuel Statique et Moteur. Ces différences inter-modalitaires peuvent être dues à des différences entre les seuils de discrimination des différents canaux sensoriels. Par exemple, du point de vue visuel, on pourrait envisager qu’on discrimine mieux deux ellipses visuelles d’excentricités proches en situation statique plutôt que dynamique. Cependant, cette hypothèse reste à être vérifiée par des expériences formelles.

Concernant la comparaison entre les variabilités inter et intra-sujet, il est intéressant de constater que dans chacune des trois expériences purement sensorielles, i.e. Visuel Statique, Visuel Dynamique et Sonore, les variabilités inter-sujet et intra-sujet ont des grandeurs comparables. Dans le cas du prototype Moteur, on constate que la variabilité inter-sujet est entre deux et trois fois plus grande que la variabilité intra-sujet. Cela signifie que chaque sujet a dessiné des prototypes dans des proportions

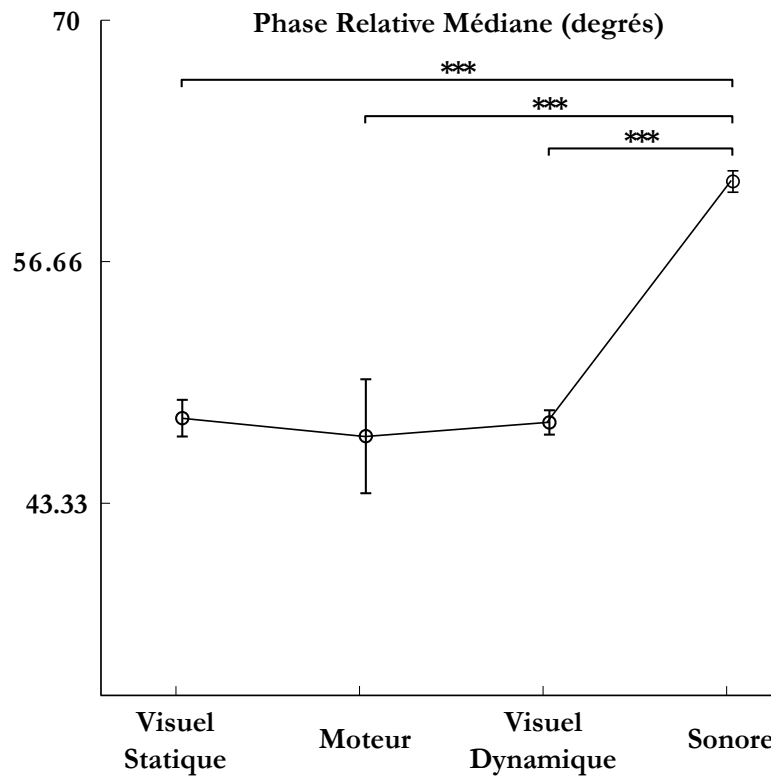


FIGURE 3.9 – Prototypes mis en évidence dans les 4 expériences. Le prototype sonore semble donc avoir une phase relative (respectivement une excentricité) supérieure (respectivement inférieure) aux trois autres prototypes.

très proches mais que la variabilité entre les sujets est-elle beaucoup plus forte. Cette différence est classique et avait déjà été observée dans les expériences de Wamain *et al.* (2011).

	Visuel Statique	Moteur	Visuel Dynamique	Sonore
Inter-sujet	$7.28 \pm .24$	9.27 ± 1.03	$12.71 \pm .48$	$11.03 \pm .97$
Intra-sujet	$6.71 \pm .23$	$3.77 \pm .42$	$12.39 \pm .39$	$10.66 \pm .62$

TABLE 3.8 – Valeurs moyennes et écarts-type des JND à travers et entre les sujets

3.6.3 Conclusion intermédiaire

On peut conclure que le prototype Sonore est significativement plus arrondi que les trois autres prototypes. Afin d'expliquer cette différence on peut formuler deux hypothèses :

1. soit le prototype Sonore est effectivement par nature plus arrondi que les trois autres nécessitant de revoir nos hypothèses initiales sur l'origine dynamique des prototypes, auquel cas on devrait conclure que, bien que prototypique, la dynamique n'explique pas à elle seule la géométrie du prototype sonore et que l'expérience visuelle prédomine dans l'émergence de la géométrie du prototype.
2. une autre hypothèse, qui ne va pas à l'encontre de notre hypothèse initiale, est que le prototype Sonore qui a été ajusté peut quand même correspondre aux trois autres prototypes, mais dans la mesure où la dynamique sonore n'est pas nécessairement intégrée de la même façon que la dynamique visuelle, il n'y a pas de relation strictement bijective entre les excentricités sonores et visuelles alors qu'il y en a une entre visuel et moteur.

Afin de trancher entre ces deux hypothèses, il est nécessaire de déterminer la relation entre l'excentricité visuelle utilisée pour synthétiser le profil de vitesse puis le son et l'excentricité de l'ellipse imaginée à partir du son. Pour cela, une expérience supplémentaire va être réalisée ayant objectif de déterminer la correspondance entre l'excentricité visuelle théorique et l'excentricité de l'ellipse imaginée à partir du son.

3.7 Expérience 8 - Comparaison entre les métriques sonore et visuelle


Dans cette expérience le but est d'évaluer si la géométrie de l'ellipse imaginée à partir du son est bien cohérente avec celle de l'ellipse théorique utilisée pour générer le son. On souhaitait ainsi déterminer l'excentricité de l'ellipse imaginée par les sujets en entendant les sons. Les sujets devaient donc ajuster une ellipse visuelle statique présentée sur un écran afin que son excentricité corresponde à celle de l'ellipse que leur évoquait le son de frottement qu'ils entendaient en même temps.

3.7.1 Méthodes

Participants

20 sujets (4 femmes) d'un âge moyen de 30.65 ans ($SD = 10.7$) ont pris part volontairement à ce test. 10 sujets étaient naïfs au sujet de l'expérience, tous avaient une audition normale et tous ont donné leur consentement écrit avant de prendre part au test. Parmi les 20 sujets, 10 d'entre eux ont participé soit à l'expérience visuelle dynamique, soit à l'expérience motrice, mais aucun n'a participé à l'expérience visuelle statique ou sonore. De plus, cette expérience s'est déroulée au minimum un mois après qu'ils aient effectué une des deux autres expériences.

Stimuli

 Les ellipses visuelles statiques et les sons de frottement de synthèse utilisés dans cette expérience ont été générés avec les mêmes méthodes que dans les expériences 4 et 7 respectivement. Dans cette expérience, les sujets devaient ajuster l'excentricité visuelle d'une ellipse afin qu'elle corresponde à l'ellipse sonore que les sujets entendaient. Les sons de frottement étaient donc fixés et les sujets ne pouvaient agir que sur l'excentricité de l'ellipse visuelle. Afin d'évaluer une distorsion éventuelle entre l'excentricité sonore et l'excentricité visuelle, l'excentricité sonore est échantillonnée avec 21 valeurs sur l'intervalle $[0, 1]$: {0; .1; .2; .3; .4; .5; .6; .65; .7; .725; .75; .775; .8; .825; .85; .875; .9; .925; .95; .975; 1.}, l'intervalle est volontairement échantillonné plus finement autour des valeurs proches de .9. Cela correspond aux phases relatives suivantes {0; 1.43; 2.87; 4.31; 5.75; 7.20; 8.66; 10.13; 11.61; 13.11; 14.63; 16.17; 17.72; 20.91; 24.22; 31.30; 39.20; 48.26; 58.99; 72.20; 89.23} (degrés). Comme les expériences 4 et 7 n'ont révélé aucune différence entre les facteurs expérimentaux ((taille) et (taille) \times (fréquence) respectivement), une seule condition expérimentale est évaluée dans cette expérience : $A = 5$ cm et $T = 1.5$ s. Chaque excentricité sonore a été répétée 4 fois avec 4 valeurs initiales {0; .25; .5; 1.} et le test était finalement composé de 84 essais 21 (excentricités) \times {4 répétitions}.

Appareillage

Les sujets étaient assis dans une cabine audiométrique, les sons étaient diffusés par le biais d'un casque Sennheiser HD-650 à une fréquence d'échantillonnage de 44100 Hz. Les sujets pouvaient ajuster l'ellipse visuelle statique à l'aide d'un clavier MIDI AKAI MPK. Les sujets étaient assis devant un écran DELL 1907fp de résolution 1280 \times 1024 et de fréquence de rafraîchissement 60 Hz. Les ellipses étaient affichées au centre de l'écran. L'interface de test a été programmée avec le logiciel Max/MSP.

Tâche

Dans cette expérience, les sujets devaient ajuster l'ellipse visuelle correspondant au son de frottement qu'ils entendaient en même temps. Il leur était précisé que le son de frottement était celui produit par une personne dessinant une ellipse, qui est une forme comprise entre un trait et un cercle. De plus, il leur était précisé que tout le panel d'ellipse était présent dans le test, du trait au cercle.

De la même façon que dans l'expérience 4, le protocole d'ajustement a été conçu de façon à éviter au maximum les biais non sensoriels, 6 touches permettaient d'ajuster les ellipses visuelles (" $< < <$ "; " $< <$ "; " $<$ " et " $>$ "; " $> >$ "; " $> > >$ ") et les pas utilisés étaient les mêmes que dans l'expérience 4 (i.e. $\{-0.1; -0.01; -0.005\}$ et $\{0.01; 0.02; 0.15\}$). Le rôle des flèches a été contrebalancé entre les sujets et les différentes ellipses sonores ont été présentées selon deux séries pseudo-aléatoires, et également contrebalancées entre les sujets.

Analyse des données

Pour chaque excentricité sonore, 4 excentricités visuelles ont été collectées puis transformées en phase relative afin d'être moyennées. Ces valeurs ont finalement été moyennées entre les sujets. Pour chaque phase relative sonore, un t-test entre les valeurs finales des phases relatives visuelles calibrées et des phases relatives sonores associées a été réalisé, le risque est fixé à $\alpha = .05$ et ajusté avec les corrections de Bonferroni. Les résultats sont présentés en phase relative et en excentricité. La distorsion entre les facteurs est modélisée par une fonction polynomiale d'ordre 3.

3.7.2 Résultats

La table 3.9 présente les valeurs moyennes des phases relatives visuelles calibrées par rapport à chaque excentricité et phase relative sonore. Les figures 3.10 et 3.11 présentent les courbes de distorsion entre les phases relatives (respectivement les excentricités) visuelles théoriques les phases relatives (respectivement les excentricités) imaginées à partir du son.

Les résultats mettent clairement en évidence une distorsion entre l'excentricité de l'ellipse sonore et l'ellipse visuelle, cette distorsion est confirmée par les tests statistiques qui montrent que l'excentricité et la phase relative sonore sont significativement différentes de l'excentricité et la phase relative visuelle pour 19 points de mesures sur 21. La phase relative est globalement sous-estimée sur tout l'intervalle $[0\ 90]$ degrés. En particulier, il est notable que l'excentricité visuelle est sous-estimée par rapport à l'excentricité sonore pour les excentricités inférieures à .5 et qu'elle est surestimée pour les valeurs comprises entre .5 et 1. Le polynôme d'ordre 3 ajusté par rapport aux résultats sur les phases relatives est le suivant :

$$\phi_v(\phi_s) = 6.3721 - 0.2071\phi_s + 0.0186\phi_s^2 - 0.0001\phi_s^3 \quad (3.10)$$

où ϕ_v et ϕ_s correspondent à la phase relative visuelle et sonore respectivement⁵.

Il est désormais intéressant de reconsidérer les résultats de l'expérience 7 au regard de cette distorsion entre ellipse visuelle et sonore. En utilisant la fonction de distorsion ajustée, on peut désormais refaire les analyses statistiques faites en section 3.6 en appliquant la fonction de distorsion aux données brutes de l'expérience sonore (i.e. expérience 7). La table 3.10 présente les valeurs moyennes des prototypes réajustés avec la fonction de distorsion (phases relatives médianes moyennées). L'analyse statistique met désormais en évidence qu'aucun prototype ne diffère significativement l'un de

5. On peut obtenir une relation entre les excentricités visuelles et sonores e_v et e_s en appliquant la formule 3.3 à cette formule.

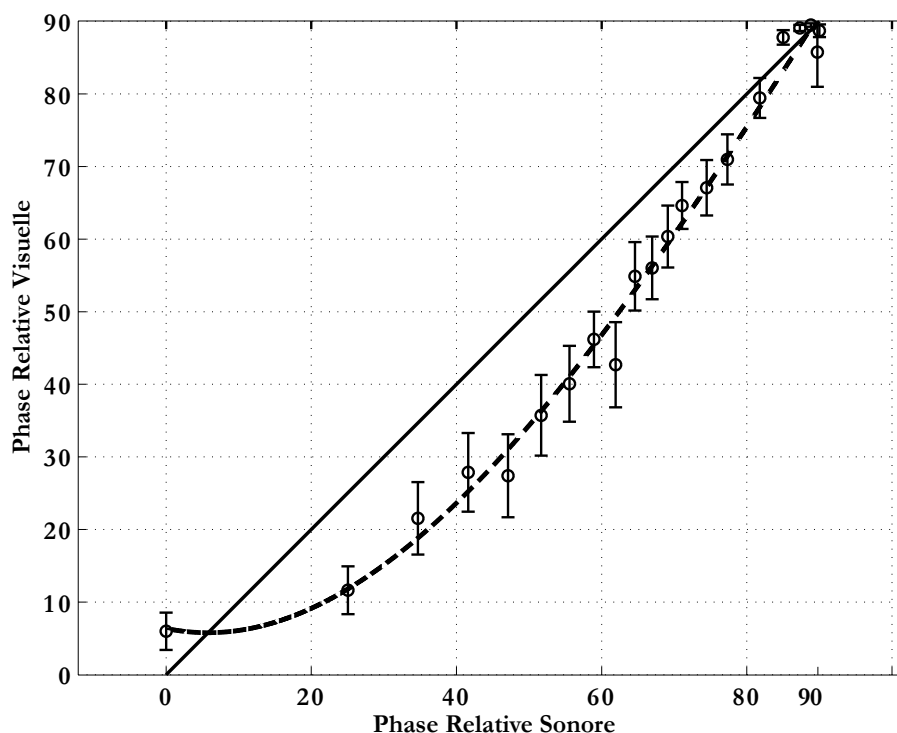


FIGURE 3.10 – Distorsion entre la phase relative sonore et la phase relative visuelle. La droite en trait plein correspond à la bissectrice et la courbe en traits discontinus au polynôme d'ordre 3 ajusté par rapport aux résultats. Chaque point de mesure correspond à la moyenne et les barres d'erreur correspondent aux erreurs standards associées.

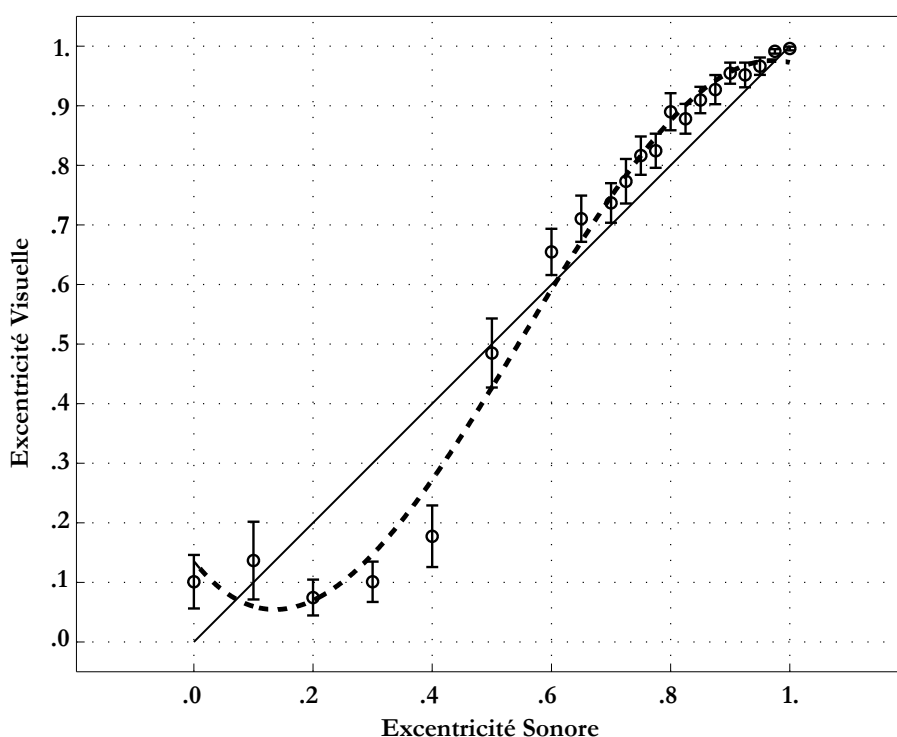


FIGURE 3.11 – Distorsion entre l'excentricité sonore et l'excentricité visuelle. La droite en trait plein correspond à la bissectrice et la courbe en traits discontinus au polynôme d'ordre 3 ajusté par rapport aux résultats. Chaque point de mesure correspond à la moyenne et les barres d'erreur correspondent aux erreurs standards associées.

ϕ_s	ϕ_v	
	M	SD
90.00	88.70	.872
89.71	85.75	4.765
88.83***	89.49	.235
87.30***	89.09	.432
85.01***	87.79	1.005
81.79	79.44	2.758
77.32***	7.98	3.458
74.46***	67.07	3.812
71.06***	64.63	3.232
69.11***	6.36	4.270
66.96***	56.04	4.317
64.58*	54.88	4.707
61.93***	42.70	5.860
58.94***	46.19	3.831
55.56***	44.08	5.228
51.67**	35.73	5.562
47.10***	27.40	5.715
41.61***	27.88	5.420
34.68***	21.55	4.991
25.06***	11.63	3.292
.00***	5.98	2.564

TABLE 3.9 – Moyennes et écarts-types de la phase relative visuelle calibrée ϕ_v en fonction de la phase relative sonore ϕ_s

l'autre ($F(3, 476) = 1.32$, $p = .27$, voir figure 3.12). On peut donc conclure que dans l'expérience sonore (expérience 7), les sujets ont bien calibré une ellipse correspondant aux ellipses prototypiques mises en évidence dans les trois autres expériences, mais qu'en raison de la distorsion entre la cinématique imaginée à partir du son et la cinématique réelle, les sujets ont réglé une excentricité plus faible (voir figure 3.13).

	Visuel Statique	Moteur	Visuel Dynamique	Sonore
ϕ_{med} (degrés)	45.17 ± 1.012 ($e_{med} = .91$)	44.52 ± 2.99 ($e_{med} = .91$)	$44.94 \pm .6723$ ($e_{med} = .91$)	$46.17 \pm .778$ ($e_{med} = .91$)

TABLE 3.10 – Valeurs moyennes et erreurs standards des phases relatives médianes des prototypes et du prototype sonore ajusté

3.7.3 Discussion intermédiaire

Cette expérience 8 avait pour objectif d'évaluer la correspondance entre une excentricité sonore donnée et l'ellipse visuelle à laquelle elle correspond du point de vue perceptif afin de tirer des conclusions des quatre premières expériences. Il s'agissait savoir si le prototype sonore de l'ellipse correspondait effectivement à une ellipse visuelle différente, ou bien si l'excentricité de l'ellipse visuelle imaginée à partir du son n'était pas cohérente avec celle utilisée pour générer le son. Les sujets devaient donc ajuster l'excentricité de l'ellipse visuelle qu'ils voyaient afin qu'elle corresponde à celle de l'ellipse sonore qu'ils entendaient. Les sons étaient composés de bruits de frottement synthétisés à partir du modèle de geste présenté en introduction. Les sujets devaient donc s'imaginer la géomé-

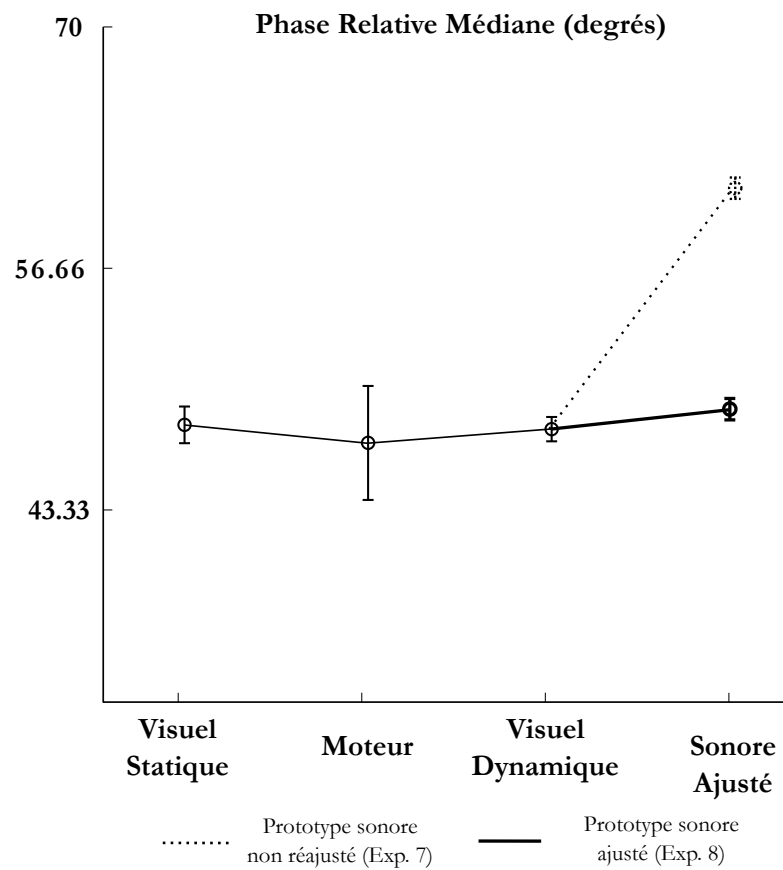


FIGURE 3.12 – Prototypes mis en évidence dans les 5 expériences. Le trait en pointillé correspond au prototype sonore non ajusté. On voit clairement que les quatre prototypes sont congruents.

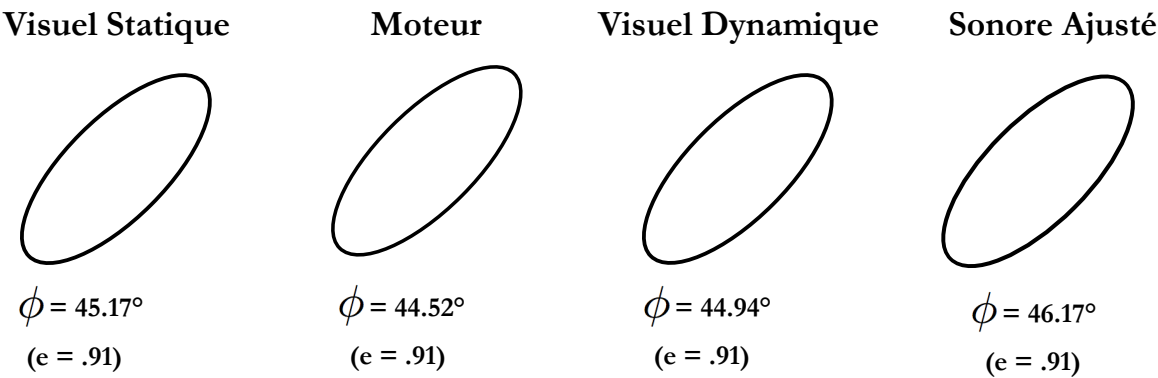


FIGURE 3.13 – Géométrie des prototypes mise en évidence dans les trois premières expériences et dans l’expérience sonore avec les résultats ajustés par rapport à la distorsion entre excentricité visuelle et sonore

trie de l'ellipse à partir du son de frottement produit par le mouvement elliptique et régler l'ellipse visuelle à partir de cette *image mentale* du geste sous-jacent au son. Les résultats ont finalement mis en évidence une distorsion évidente entre l'excentricité sonore et l'excentricité visuelle imaginée à partir du son : à une excentricité sonore donnée, l'ellipse visuelle imaginée est quasiment systématiquement plus plate. Partant de cette constatation, les résultats de l'expérience 7 ont été ré-analysés en transformant les données brutes (i.e. les excentricités calibrées) et il a été finalement mis en évidence que la géométrie du prototype mis en évidence à partir des sons de frottement était similaire à celles des trois autres prototypes.

La distorsion entre l'ellipse visuelle et l'ellipse sonore est très intéressante car elle permet de quantifier de façon très formelle la géométrie de la forme qui est imaginée à partir du son. Cette distorsion va être discutée plus en détail dans la discussion générale qui suit.

3.8 Discussion générale

Le but de cette étude était de comprendre quelle forme est réellement imaginée quand un sujet entend les sons de frottement produits par les mouvements graphiques. Pour aborder cette problématique, un problème plus général a été abordé pour évaluer dans quelle mesure les formes géométriques prototypiques à la base des processus de catégorisation et d'imagination des formes émergent de contraintes sensori-motrices, et plus particulièrement si cette géométrie particulière était la conséquence de caractéristiques dynamiques. En effet, la relation entre son, geste et forme mise en évidence dans le chapitre 2 pose nécessairement le problème de la nature géométrique de la représentation inférée par le son. Pour cela, nous nous sommes intéressés à une catégorie de formes spécifiques – les ellipses. Les ellipses sont toutes les formes comprises entre un trait et un cercle, on a ainsi cherché à vérifier si ces formes étaient caractérisées par une excentricité particulière à travers différents canaux sensoriels, mais surtout, si cette excentricité particulière était la conséquence d'une représentation basée sur la dynamique des mouvements biologiques plus que sur leurs géométries. Pour cela, une première expérience a mis en évidence que le prototype visuel statique des ellipses correspondait à une ellipse d'excentricité .91 et ne dépendait pas de la taille. Afin de comparer ce prototype visuel statique au prototype moteur, une seconde expérience a été réalisée lors de laquelle les sujets devaient cette fois dessiner en continu la forme elliptique qui était la plus représentative de cette catégorie de formes géométriques. Il leur était demandé de dessiner des ellipses de trois tailles différentes, une petite, une moyenne, et une grande. Les résultats ont mis en évidence que les sujets ont tous dessiné des ellipses dans les mêmes proportions quelle que soit la taille, et de plus que l'excentricité de cette ellipse prototypique du point de vue moteur était de l'ordre de .91, comme celle du prototype mis en évidence du point de vue visuel statique. Ces deux premières expériences ont donc suggéré que la géométrie du prototype peut émerger de contraintes sensori-motrices dans la mesure où le prototype moteur est cohérent avec le prototype visuel. Afin de montrer que le prototype visuel est avant tout dynamique, une troisième expérience a été réalisée lors de laquelle les sujets devaient calibrer l'excentricité d'une ellipse parcourue par un point lumineux afin qu'elle leur soit la plus représentative de cette catégorie de forme. En plus de la taille, la fréquence du mouvement a également été testée. Les résultats ont mis en évidence que les sujets ont calibré l'excentricité de l'ellipse autour de .91 comme dans les deux expériences précédentes, et qu'enfin, ni la taille ni la fréquence du geste n'ont influencé la valeur de l'excentricité du prototype. Ces trois premières expériences vont donc dans le sens de nos hypothèses sur le fait que l'excentricité du prototype de l'ellipse peut émerger de caractéristiques biomécaniques. Cependant dans ces trois expériences, la géométrie du prototype était soit évaluée toute seule, soit dans des cas dynamiques où la géométrie

était également disponible, et on ne peut donc pas trancher sur le fait que le prototype soit purement dynamique, et que le profil de vitesse du mouvement (du point lumineux dans l'expérience 6 ou de la main des sujets dans l'expérience 5) soit prototypique en soi. Une façon élégante d'évaluer si la dynamique biologique de cette ellipse particulière est prototypique en soi est d'utiliser la modalité sonore. En effet, la modalité visuelle ou la sortie motrice mettent en jeu à la fois la géométrie de la forme et la dynamique, et elle n'est donc pas évaluée toute seule. La modalité sonore, par sa capacité à transmettre une dynamique via des variations de timbre, a permis de mettre en place une expérience durant laquelle les sujets devaient calibrer le son de frottement produit par les mêmes mouvements que ceux de l'expérience visuelle dynamique afin que le geste évoqué par ces sons leurs évoque la trajectoire la plus représentative de la catégorie de forme ellipse. Les sujets ne voyaient pas la forme qu'ils calibraient et ne pouvaient se baser que sur le geste évoqué par le son pour imaginer une dynamique. Les résultats de cette expérience ont bien montré qu'une dynamique était prototypique, cependant, contrairement aux expériences précédentes, l'excentricité correspondante était de .8, soulevant de ce fait une question principale : est-ce que l'ellipse imaginée par les sujets quand ils ont calibré les sons correspondait à celle qui était effectivement transmise dans les variations de timbre des sons ? Afin de pouvoir trancher sur le fait que le prototype sonore correspondait ou non à celui mis en évidence dans les trois premières expériences, une cinquième expérience a été réalisée pour évaluer la correspondance perceptive entre l'ellipse imaginée à partir du son et l'ellipse visuelle correspondante. Dit autrement, cette expérience avait pour objectif de voir si l'ellipse que les sujets se sont imaginée à partir du son avait la même excentricité que celle utilisée pour générer le stimulus sonore. Les sujets devaient ainsi ajuster l'excentricité de l'ellipse visuelle statique par rapport à celle qu'ils imaginaient en entendant le son de frottement. Cette expérience a mis en évidence une distorsion entre l'excentricité de l'ellipse sonore et celle de l'ellipse visuelle statique correspondante, l'excentricité de l'ellipse imaginée était systématiquement surestimée par rapport à celle réellement transmise dans le son. De plus, en transformant les résultats de l'expérience 7 à partir de cette distorsion, on a pu montrer qu'en fait, l'ellipse sonore prototypique que les sujets avaient calibrée dans l'expérience sonore correspondait bien aux trois prototypes mis en évidence dans les trois premières expériences. Cette dernière expérience nous oblige à discuter l'ensemble des résultats selon deux aspects. Tout d'abord le fait que lorsqu'on entend le son de frottement produit par le mouvement d'un point respectant les mouvements biologiques, la forme imaginée visuellement ne correspond pas à celle qui est initialement transmise dans le son. Dans un second temps, ces résultats donnent un éclairage nouveau sur la notion de prototype géométrique et la façon dont ils peuvent émerger des caractéristiques biomécaniques dynamiques et de notre expérience sensori-motrice.

L'expérience 8 a donc montré que la forme imaginée à partir du mouvement évoqué par les sons de frottement n'était pas cohérente avec la forme géométrique associée. Afin de tenter de comprendre cette distorsion, il est nécessaire de se focaliser sur les informations transmises par les stimuli sonores. En effet, les sons de frottement de synthèse ont été générés à partir du profil de vitesse du mouvement respectant la loi en puissance $1/3$ et en associant ces variations à une variation de fréquence centrale. Cette expérience nous renseigne donc de façon quantitative sur la façon dont on intègre l'information cinématique à partir des variations de timbre. Dans le chapitre 2, on a vu que l'on était capable de reconnaître les variations de timbre associées à la loi en puissance $1/3$, et que de plus, on pouvait discriminer des formes géométriques sur la base des sons de frottement. Ces expériences avaient donc mis en évidence la pertinence de ces informations, i.e. *invariantes* dans une perspective écologique, dans la relation entre les sons et les gestes. En ne s'intéressant qu'à une seule forme, et dans le cadre d'une dynamique qui respecte la loi en puissance $1/3$, on peut voir ici de façon évi-

dente que la dynamique du mouvement évoquée par le son, bien que caractéristique, n'est pas pour autant intégrée en cohérence totale avec celle introduite dans le stimulus de synthèse. Le fait que l'excentricité de la forme imaginée à partir du son soit quasi systématiquement inférieure à celle de l'ellipse visuelle correspondante montre que le profil de vitesse imaginé n'est pas celui initialement transmis et en particulier qu'il est surestimé. En effet, à une ellipse d'excentricité .8 correspond en fait une ellipse d'excentricité .9, les variations de vitesse du mouvement associées à la première sont donc plus petites que celles associées à la seconde. Le mapping entre la vitesse et la fréquence et la vitesse du mouvement a été choisi comme étant linéaire, i.e. $f(t) \propto v(t)$, mais finalement on se rend compte que sa perception n'est pas linéaire, i.e. $f(t) = d(v(t))$ où d n'est pas une fonction linéaire. La distorsion d peut avoir une première origine purement psychoacoustique, en effet, il est bien connu que la perception de la fréquence ne respecte pas une échelle linéaire et qu'elle se modélise bien par une échelle logarithmique (Moore et Glasberg, 1983). La représentation des variations de fréquences des stimuli transmise par le système auditif n'est donc pas bijective avec celle introduite dans les sons de frottement. Une hypothèse possible est donc que cette distorsion a faussé la représentation sous-jacente induite par le son. Cependant, bien qu'une telle distorsion structurelle existe, les processus cognitifs devraient l'avoir intégrée et la représentation sous-jacente ne devrait être impactée pour autant.

Le problème de s'imaginer une forme à partir d'un son monophonique est un problème compliqué, en effet, quand une forme est définie spatialement par deux dimensions, les variations de timbre n'en transmettent qu'une. La tâche est donc complexe, et les sujets doivent s'imaginer la forme qui correspond de façon la plus plausible au son à partir d'une seule dimension cinématique. Cette tâche ne peut se faire que sur la base d'une inférence basée sur un modèle de relation entre une variation de vitesse (ici intégrée par le biais d'une variation de timbre) et les formes qui peuvent y correspondre. Dans le chapitre 2, on a fait l'hypothèse que c'était la loi en puissance $1/3$ qui permettait de remonter aux formes à partir de la dynamique et de les discriminer en reconnaissant les phases d'accélération et de ralentissement. Ici le problème est plus complexe car les mouvements elliptiques ont des variations cinématiques périodiques et l'on veut vraiment évaluer, pour une variation de vitesse donnée, quelle est la géométrie de la forme qui correspond le plus au son entendu.

Du point de vue visuel, la perception des mouvements biologiques a été largement étudiée et de nombreuses expériences ont montré que lorsque l'on perçoit un point lumineux qui se déplace sur une forme en respectant la loi en puissance $1/3$, les variations de vitesse ne sont pas perçues et on perçoit un mouvement globalement uniforme (Viviani et Stucchi, 1992; Bidet-Ildei *et al.*, 2006; Levitt-Binnun *et al.*, 2006). Différentes hypothèses ont été proposées afin d'expliquer pourquoi la perception et la production des mouvements biologiques partagent des caractéristiques communes. Une hypothèse intéressante dans la perspective de la discussion de nos résultats provient de considérations mathématiques sur l'espace géométrique (la métrique) dans lequel les mouvements sont considérés. Il a été montré que si l'on considère les mouvements graphiques non plus dans un espace euclidien mais dans un espace affine, une vitesse affine nulle correspond à une vitesse euclidienne respectant la loi en puissance $1/3$ (Pollick et Sapiro, 1997; Flash et Handzel, 2007). De plus dans un espace affine, les grandeurs sont invariantes à une dilatation spatiale près, ce qui est très intéressant du point de vue fonctionnel. Cela permet en effet au système perceptif de s'affranchir de l'orientation des yeux et de la distance à la forme (Pollick et Sapiro, 1997). Dans l'hypothèse où l'apprentissage des gestes se réalise en cohérence avec les informations visuelles, il semble naturel que nos mouvements pos-

sèdent également ces caractéristiques⁶. Au regard de cette hypothèse d'intégration visuelle basée sur une métrique affine des informations cinématiques, le prototype dynamique devra respecter la loi en puissance $1/3$ et correspondre ainsi à un mouvement uniforme. À l'inverse, quand on entend le son de frottement produit lors du dessin de tels mouvements respectant les mouvements biologiques, les variations de vitesse sont clairement perçues (Thoret *et al.*, 2014c) et le son transmet la cinématique dans sa dimension euclidienne. Dit plus simplement, quand la vision lisse les variations de vitesse pour les percevoir de façon uniforme, l'audition les transmet dans toutes leurs subtilités. Dans l'expérience de calibration sonore, les sujets ont donc essayé de faire correspondre la dynamique euclidienne évoquée par le son à celle qu'ils incarnent par le biais du système visuel, basé lui sur une métrique affine. La non correspondance entre la métrique affine et la métrique euclidienne peut donc expliquer pourquoi le prototype calibré du point de vue sonore présente une excentricité plus faible que dans les trois autres expériences. En essayant de faire correspondre la cinématique évoquée par le son à celle correspondant au prototype, les variations évoquées par les sons correspondant aux excentricités proches de .9 semblent avoir été trop importantes ce qui a amené les sujets à diminuer l'excentricité de la forme pour obtenir un *taux de variation* de la vitesse acceptable par rapport à leur attente. Cela met donc en évidence que l'invariant transformationnel cinématique extrait par le système auditif est interprété à partir de considérations pouvant également émaner de la modalité visuelle. Sans remettre en cause le fait que le profil de vitesse transmis par les sons de frottement fournit une information pertinente sur la cinématique du *vrai* geste, cette expérience montre que la relation entre les sons de frottement produits dans un contexte écologique et leur *représentation* en terme de forme géométrique n'est pas directe. Notre expérience ne permet cependant pas de trancher entre ces deux hypothèses et des expériences complémentaires sont nécessaires afin de bien comprendre l'influence de chaque modalité dans la représentation sous-jacente de la dynamique. Cependant, cette dernière expérience nous a permis de mettre en évidence que le prototype calibré dans l'expérience 7 correspondait bien à celui des trois premières expériences étant donnée la distorsion entre les excentricités sonores et les excentricités visuelles imaginées à partir des sons, on peut donc quand même discuter la question de l'origine de ces prototypes.

Notre hypothèse initiale était que les caractéristiques spatiales des prototypes géométriques pouvaient avoir une origine issue de contraintes biomécaniques, principalement dynamiques, modelant ainsi nos représentations par le biais de notre expérience sensori-motrice. Cette hypothèse d'*inscription corporelle de l'esprit* semble de plus en plus évidente (Varela *et al.*, 1993; Barsalou, 1999), la réactivation de schémas moteurs à la présentation concrète d'un objet a été mise en évidence dans plusieurs études (Chao et Martin, 2000; Creem-Regehr et Lee, 2005; Grafton *et al.*, 1997; Grèzes et Decety, 2002) et en particulier dans le cas des formes géométriques (Longcamp *et al.*, 2003; James et Gauthier, 2006; Longcamp *et al.*, 2006). Cette réactivation aide au développement cognitif et souligne l'extrême importance de la perception par l'action pour le développement de comportements adaptés. Pour ce qui est des formes géométriques, le groupe d'Edouard Gentaz s'est intéressé dans plusieurs études à essayer de comprendre l'origine des prototypes géométriques. Kalenine *et al.* (2011) ont mis en évidence l'importance de l'exploration haptique pour la reconnaissance des formes géométriques. Kalénine *et al.* (2013) ont montré que des enfants de 5 ans dessinaient spontanément des rectangles

6. L'origine des caractéristiques communes entre perception et production des mouvements biologiques est encore débattue et nous n'avons présenté ici que le point de vue où la loi en puissance $1/3$ est la conséquence d'un traitement au niveau du système nerveux central. Un autre point de vue également pertinent propose qu'au niveau moteur, la loi en puissance $1/3$ est uniquement conséquence de caractéristiques biomécaniques, et en particulier du couplage non linéaire entre deux oscillateurs harmoniques. Cette hypothèse postule que cette relation a une origine *périphérique* et que les processus perceptifs en découlent.

autour d'un prototype géométrique commun de ratio longueur sur largeur de 2.3. Theurel *et al.* (2012) ont quant à eux montré qu'il y avait pas d'effet de prototype dans la reconnaissance haptique des formes en comparant les résultats de tâches de reconnaissance haptique de formes géométriques chez des adolescents aveugles de naissance et des sujets avec les yeux bandés. Finalement, leur hypothèse est que les prototypes géométriques émergent principalement d'une expérience visuelle : (a) les prototypes sont les formes partageant le plus de traits communs avec tous les autres membres de la catégorie, (b) les prototypes sont les formes de la catégorie les plus distincts des formes des autres catégories. Cela avait déjà été suggéré dans les études séminales d'Eleanor Rosch (Rosch, 1973, 1975; Rosch et Mervis, 1975). Sans rejeter ces conclusions et l'importance de l'expérience visuelle dans l'émergence des prototypes, nous avons ici montré de façon évidente l'importance des contraintes biomécaniques, portant intrinsèquement des caractéristiques dynamiques associées à la production de formes qui ensuite deviennent prototypiques d'un point de vue géométrique, dans la formation des caractéristiques géométriques du prototype. Wamain *et al.* (2011) avaient déjà suggéré cela dans deux expériences qui avaient mis en évidence que des sujets adultes discriminaient mieux visuellement les ellipses d'excentricité qu'ils dessinent préférentiellement. Grâce à l'expérience sonore, nous avons pu montrer que la dynamique du prototype était effectivement prototypique en soi indépendamment de la géométrie, et c'est à notre connaissance la première étude le montrant de cette façon. De plus, la valeur de l'excentricité du prototype étant cohérente à travers les différentes expériences⁷, cela montre qu'il y a vraiment une correspondance sensori-motrice (entre la géométrie visuelle et la dynamique motrice). De plus, la valeur de l'excentricité prototypique mise en évidence est critique car elle est cohérente avec la valeur de l'ellipse préférentielle mise en évidence dans des études purement biomécaniques (Dounskaia *et al.*, 2000; Sallagoity *et al.*, 2004; Danna *et al.*, 2011).

Par ailleurs, Kalénine *et al.* (2013) ont montré que chez les adultes, les rectangles prototypiques avaient un ratio longueur/largeur de 2.3 (SD=.79). Pour les ellipses nous avons jusqu'à présent uniquement considéré l'excentricité, mais si on souhaite comparer les proportions de l'ellipse prototypique que l'on a mise en évidence à celles des rectangles prototypiques, on peut également regarder le ratio entre son demi-grand axe et son demi petit axe⁸. On peut alors remarquer que l'ellipse prototypique correspond à un ratio de 2.29 similaire à celui du rectangle prototypique mis en évidence par Kalénine *et al.* (2013), voir figure 3.14. Bien que la coordination motrice nécessaire à la production



FIGURE 3.14 – L'ellipse prototypique mise en évidence dans cette étude a des proportions similaires à celles du rectangle prototypique mis en évidence dans une tâche de production par Kalénine *et al.* (2013)

d'un rectangle diffère de la production d'une ellipse, le rectangle est également une forme fermée et on peut faire l'hypothèse qu'une telle coordination est régie par des règles similaires sujettes à l'existence d'attracteurs du mouvement mettant en évidence des aspects dynamiques préférentiels comme dans le cas des ellipses (Dounskaia *et al.*, 2000). En effet, on peut envisager le dessin d'un rectangle comme la coordination entre 4 phases relatives successives, et il serait intéressant de voir si la géométrie du prototype rectangulaire mise en évidence par Kalénine *et al.* (2013) ne peut pas également émerger d'une modélisation basée sur un système d'oscillateurs couplés. Ces résultats ne viennent

7. On considérera à partir d'ici que le prototype sonore a également une excentricité de .91 en raison des résultats de l'expérience 8.

8. Le ratio R entre le demi grand axe et le demi petit axe se calcule très facilement à partir de l'excentricité : $R = \sqrt{1 - e^2}$

pas contredire ceux déjà mis en évidence mais ils montrent qu'il est nécessaire d'aborder le problème de l'origine des prototypes en replaçant la motricité au centre des débats et non comme unique sortie pour évaluer la géométrie du prototype. De plus, l'utilisation de la modalité auditive pour investiguer la représentation des formes géométriques ouvre de nombreuses perspectives pour considérer séparément la dynamique dans le processus de formation des prototypes.

Du point de vue de la synthèse sonore, ces expériences, et plus particulièrement l'expérience 8, permettent également d'envisager de calibrer précisément le mapping en la vitesse et la fréquence centrale du filtre passe-bande utilisé dans le modèle de synthèse afin de générer des sons cohérents avec un profil de vitesse donné et une forme géométrique particulière. Ces aspects seront plus détaillés dans la conclusion générale (Chapitre 6).

Chapitre 4

Voir des cercles, dessiner des ellipses : quand l’audition trompe la perception visuelle des mouvements biologiques

Sommaire

4.1	Introduction	104
4.2	Méthodes générales	107
4.3	Expérience préliminaire – Couplage audio-moteur	113
4.4	Expérience 9 – Couplage visuo-moteur	114
4.5	Expérience 10 – Couplage audio-visuo-moteur	117
4.6	Discussion générale	121

Dans le chapitre précédent, nous avons vu que lorsque nous nous imaginons une ellipse dessinée à partir d’un son de frottement, la forme imaginée était beaucoup plus arrondie que la forme réellement dessinée. Une hypothèse possible est donc que les métriques visuelle et sonore intègrent de façon différente les informations cinématiques, pour une même cinématique l’association du mouvement à une forme géométrique est donc différente selon la modalité. Cependant, le monde sensible ne fait rarement intervenir qu’une seule modalité et nous percevons le monde par le biais de tous nos sens simultanément. Ainsi, il semble intéressant de se demander quelle va être la forme perçue quand des informations cinématiques différentes sont perçues simultanément par le système visuel et le système auditif. De plus, on a vu dans l’état de l’art que la perception visuelle des mouvements était naturellement biaisée par les contraintes motrices et notamment la loi en puissance $1/3$. Les sons évoquant des mouvements comme les sons de frottement générés quand une personne dessine semblent donc être des candidats sérieux pour altérer voir modifier la perception visuelle et la reproduction d’un mouvement visuel. Ce chapitre présente ainsi des expériences de synchronisation sensori-motrice avec des stimuli, tout d’abord sonores, puis visuels, et enfin audiovisuel. Afin d’évaluer dans quelle mesure un stimulus sonore ne peut pas interférer significativement avec la perception visuelle et sur la reproduction graphique d’un mouvement visuel.

4.1 Introduction

Percevoir et agir de façon cohérente dans l’environnement nécessite l’intégration d’informations émanant de nos différents canaux sensoriels de façon synchrone. La perception du mouvement est

une compétence essentielle qui assure une cohérence spatiotemporelle entre nos actions et leurs conséquences sur l'environnement (Gibson, 1966; Shepard, 1984; Maravita *et al.*, 2003). De nombreuses études ont mis en évidence que les mouvements biologiques impliquaient des cinématiques très spécifiques, et en particulier la loi en puissance $1/3$ reliant la vitesse d'un mouvement à la courbure de sa trajectoire (Lacquaniti *et al.*, 1983). De plus les capacités humaines à percevoir et reproduire un mouvement sont meilleures quand les mouvements cibles respectent les règles des mouvements biologiques (Johansson, 1973) et en particulier la loi en puissance $1/3$ (Viviani *et al.*, 1987). Quand la dynamique visuelle d'un point lumineux ne respecte pas les règles des mouvements biologiques, la perception de la géométrie de la trajectoire est modifiée et la géométrie perçue ne correspond plus à celle réellement parcourue (Viviani et Stucchi, 1989). De cette façon, une trajectoire circulaire parcourue avec un profil de vitesse correspondant à une trajectoire elliptique, i.e. qui accélère dans les deux parties opposées du cercle, est perçue comme plus elliptique que si le point se déplace à vitesse constante sur le cercle. Les mêmes résultats ont été mis en évidence du point de vue kinesthésique dans une expérience où les sujets devaient juger de la géométrie de la trajectoire parcourue par leur main qui était déplacée par un bras articulé avec différentes cinématiques (Viviani *et al.*, 1997). À l'inverse, la vitesse d'un point lumineux se déplaçant sur une trajectoire elliptique à vitesse constante n'est pas perçue comme telle, mais comme accélérant dans les parties les plus courbées de la trajectoire (Viviani et Stucchi, 1992). Ces résultats suggèrent que les sujets basent leurs jugements visuels sur la covariation entre vitesse et courbure, et quand cette covariation ne respecte pas celle des mouvements biologiques, la perception de la dynamique et de la trajectoire parcourue est alors modifiée. En plus de ces résultats purement perceptifs, une étude de Viviani a montré que le tracking visuo-moteur, c'est-à-dire la capacité à suivre un point lumineux se déplaçant chaotiquement en le pointant avec un stylo par exemple, était la meilleure quand la dynamique du point lumineux respectait la loi en puissance $1/3$ (Viviani *et al.*, 1987). Tous ces résultats suggèrent que les mouvements biologiques sont intégrés à un niveau amodal ne relevant *a priori* ni de la modalité visuelle, ni de la modalité kinesthésique (Viviani *et al.*, 1997).

Malgré le grand intérêt suscité dans la communauté scientifique pour la perception des mouvements biologiques, peu d'études ont à notre connaissance abordé la perception du mouvement biologique du point de vue de la modalité auditive. Pourtant, il est bien connu que les sons révèlent naturellement une dynamique à travers leurs variations intrinsèques de timbres. Nos réactions gestuelles et corporelles à un rythme perçu sont même plus importantes pour une stimulation sonore que visuelle (Repp et Penel, 2004). On peut ainsi très facilement générer des sons qui évoquent des mouvements se basant sur des morphologies sonores spécifiques. Par exemple en faisant varier la brillance ou le centre de gravité spectral d'un son au cours du temps, il a été mis en évidence que l'on pouvait évoquer des classes particulières de mouvements (Chowning, 1971; Kronland-Martinet et Voinier, 2008; Merer *et al.*, 2013). Concernant, les mouvements biologiques, une étude récente a mis en évidence que l'on pouvait reconnaître la loi en puissance $1/3$ juste en se basant sur les variations de timbre du son de frottement produit par une personne qui dessine (Thoret *et al.*, 2014c) (cf. Chapitre 2). En reliant les caractéristiques cinématiques des mouvements graphiques aux variations de timbre d'un son de frottement de synthèse, cette étude a également montré que l'évocation sonore de la cinématique d'un geste suffisait à discriminer des formes géométriques simples.

En comparant ces derniers résultats à ceux concernant la modalité visuelle, il est intéressant de constater que l'audition et la vision ne fournissent pas la même information sur la dynamique. En effet, les résultats du chapitre précédent ont mis en évidence que les métriques visuelles et sonores ne concordaient pas dans la perception de la dynamique. De plus, si les variations de vitesse sont

bien perçues du point de vue sonore quand les mouvements respectent la loi en puissance $1/3$, ces variations ne sont plus perçues et la vitesse du mouvement est alors perçue comme constante (Viviani et Stucchi, 1992). Ainsi, on peut faire l'hypothèse que la perception visuelle d'un mouvement peut être complètement modifiée en présence d'un son révélant clairement une autre dynamique que celle du stimulus visuel. Dans l'étude qui suit, on fait l'hypothèse qu'ajouter une dynamique sonore à un stimulus visuel peut complètement distordre la perception visuelle de la géométrie de la forme parcourue par le point lumineux, de la même façon qu'une dynamique non biologique peut distordre la forme perçue dans le cas purement visuel. De plus, on fait l'hypothèse que si la dynamique visuelle est déjà non-biologique, par exemple celle d'un cercle parcouru avec des accélérations dans les parties opposées, rajouter une dynamique sonore congruente avec la dynamique visuelle non-biologique va alors amplifier la distortion visuelle. Il a en effet été mis en évidence que la corrélation d'informations émanant de modalités différentes améliorerait l'intégration de l'information, par exemple si un *bip* sonore est présenté dans une position congruente avec un flash lumineux, notre capacité à localiser la position du flash lumineux est alors améliorée (Parise *et al.*, 2012). Ainsi, en analysant les combinaisons de dynamiques visuelles et sonores congruentes ou non, il semble possible d'évaluer le poids de chaque modalité dans les processus de perception et d'action, processus connus pour être régis par des processus multimodaux (Zmigrod et Hommel, 2013).

Dans toutes les situations qui ont été étudiées, afin de déterminer la forme perçue, nous avons fait l'hypothèse qu'en demandant à des sujets de synchroniser leurs gestes avec un mouvement audiovisuel en boucle ouverte visuelle (i.e. sans voir leur main), nous pourrions déterminer la géométrie de la forme perçue et ainsi évaluer l'influence du son sur la perception visuelle des caractéristiques géométriques de chaque situation. Les caractéristiques géométriques et cinématiques des performances motrices des sujets doivent alors révéler si les sons ont altéré la perception visuelle dans les différentes situations audiovisuelles. Les formes géométriques considérées étaient soit des ellipses soit des cercles, et les dynamiques visuelles respectaient ou non les caractéristiques des mouvements biologiques, i.e. la loi en puissance $1/3$.

Une telle tâche implique donc la production de mouvements elliptiques répétitifs, et il a été mis en évidence que ce type de mouvement impliquaient des états biomécaniques stables (Dounskaia *et al.*, 2000), aussi appelés attracteurs moteurs déjà présentés dans le chapitre précédent. Les performances motrices devront donc être analysées au regard de ces connaissances sur les caractéristiques biomécaniques de ces mouvements afin de voir dans quelle mesure les dynamiques visuelles non-biologiques et l'ajout de sons évoquant une dynamique ont également pu modifier ces attracteurs.

Deux expériences ont ainsi été menées. La première expérience était purement visuelle et avait pour objectif de vérifier que les caractéristiques géométrico-dynamiques d'un mouvement visuel biaisaient effectivement sa perception et sa reproduction en boucle ouverte dans le cas où elles ne respectaient pas les règles des mouvements biologiques. Cela a déjà été mis en évidence par Viviani et ses collègues dans des tâches purement psychophysiques (Viviani et Stucchi, 1989), et dans des tâches de reproduction en boucle fermée (Viviani *et al.*, 1997), i.e. tracking d'une cible avec vision de la main, mais jamais dans une tâche en boucle ouverte visuelle. Ces résultats devraient donc compléter les connaissances sur la perception et la reproduction des mouvements biologiques visuels. Dans la seconde expérience, les sujets devaient synchroniser leurs mouvements avec des stimuli audiovisuels, produits à partir de sons de friction de synthèse associés aux stimuli visuels précédents dans le but d'évaluer dans quelle mesure le son peut modifier significativement la perception visuelle et la reproduction d'un mouvement. Comme les performances motrices des sujets

ont été considérées comme la traduction de la forme visuelle perçue, les résultats seront discutés en tenant compte des synergies et caractéristiques biomécaniques sous-jacentes qu'impliquent ces mouvements (i.e. attracteurs moteurs).

4.2 Méthodes générales

Les deux expériences de cette étude ont des protocoles expérimentaux relativement proches, cette partie présente les méthodes communes aux deux expériences, les détails spécifiques à chacune d'entre elles seront décrites au sein de chaque expérience.


4.2.1 Participants

17 sujets (2 femmes) ont pris part aux deux expériences. Tous les sujets étaient droitiers, avaient une vision normale ou corrigée et ne présentaient aucun problème auditif. Ils étaient âgés en moyenne de 28.53 ans ($SD=7.99$) et étaient tous naïfs au sujet de l'expérience.

4.2.2 Stimuli

Les stimuli visuels sont communs aux deux expériences et sont définis dans cette section. La génération des stimuli sonores de l'expérience audiovisuelle découle directement de la méthode utilisée pour générer les stimuli visuels et sera présentée dans la seconde partie de cette section.

Définition générale des stimuli visuels.

 Les stimuli visuels sont des points lumineux blancs de 6 mm de diamètre affichés sur un écran sur un fond noir. Le mouvement du point décrit est soit une trajectoire circulaire, soit une trajectoire elliptique. Sa cinématique respecte ou non les mouvements biologiques. Un mouvement visuel est donc défini par une configuration géométrique et une cinématique visuelle (notée VK dans la suite). La configuration géométrique correspond à la géométrie de la trajectoire parcourue (i.e. cercle ou ellipse), tandis que la cinématique visuelle correspond à la façon dont la trajectoire est parcourue, i.e. dans quelles parties de la trajectoire le point lumineux accélère. Deux configurations géométriques ont été considérées :

1. un cercle, i.e. une ellipse d'excentricité $e_g = 0$, de rayon 6.36cm et de périmètre 40 cm ;
2. une ellipse horizontale d'excentricité $e_g = .9$ avec un demi grand axe de 9.05 cm et un périmètre de 43.8 cm ;

Trois configurations cinématiques sont définies à partir de différentes vitesses tangentielles qui respectent la cinématique visuelle de 3 excentricités différentes, e_{vk} à savoir : $e_{vk} = 0$ (vitesse constante), $e_{vk} = .9$ (variations de vitesse cohérentes avec celles de l'ellipse géométrique horizontale), $e_{vk} = -.9$ (variations de vitesse correspondant à une ellipse orientée orthogonalement à l'ellipse horizontale, i.e. verticalement). Le mouvement du point lumineux le long de l'ellipse cinématique est défini par ses coordonnées $(x_k(t), y_k(t))$ avec un périmètre égal à celui de la configuration géométrique associée. La méthode pour calculer $(x_g(t), y_g(t))$ en fonction de $(x_k(t), y_k(t))$ a été proposée par Viviani *et al.* (1997) et est détaillée dans le paragraphe suivant. La figure 4.1 explicite également la méthode utilisée.

Le point lumineux se déplace donc le long d'une ellipse d'excentricité e_g avec différents profils de vitesse qui sont définis par le mouvement d'un point le long d'une autre ellipse d'excentricité e_{vk} .

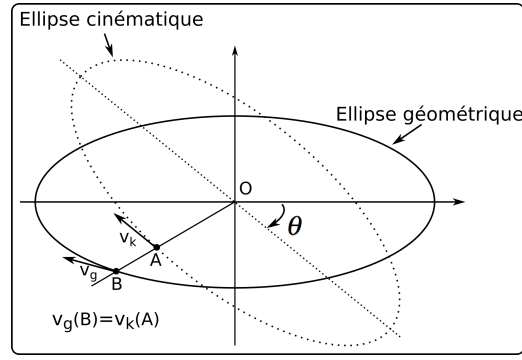


FIGURE 4.1 – Principe de générations des mouvements visuels

Le mouvement du point le long de l'ellipse géométrique e_g est défini par le mouvement de Lissajous suivant :

$$\begin{cases} x(t) = A_{xg} \sin(\phi(t)) \\ y(t) = A_{yg} \cos(\phi(t)) \end{cases} \quad (4.1)$$

où ϕ est une fonction croissante du temps telle que $\phi(0) = 0$ et $\phi(T) = T$ avec T la période d'un cycle. L'excentricité e_g de l'ellipse correspondante est alors définie par $e_g = \sqrt{1 - \left(\frac{A_{yg}}{A_{xg}}\right)^2}$ et le périmètre P_g est défini par $P_g = 4A_{xg}\Sigma(e, \pi)$ avec Σ l'intégrale elliptique incomplète. Il est important de noter qu'un mouvement paramétré de cette façon respecte la loi en puissance 1/3 (la démonstration est présentée en annexe 7.1). Par ailleurs, la trajectoire $(x_k(t), y_k(t))$ du mouvement sur l'ellipse cinématique est définie par un mouvement de Lissajous pivoté d'un angle θ par rapport à l'horizontale, voir figure 4.1 :

$$\begin{cases} x_k(t) = A_{xk} \sin\left(\frac{2\pi t}{T}\right) \cos \theta - A_{yk} \cos\left(\frac{2\pi t}{T}\right) \sin \theta \\ y_k(t) = A_{yk} \sin\left(\frac{2\pi t}{T}\right) \sin \theta + A_{xk} \cos\left(\frac{2\pi t}{T}\right) \cos \theta \end{cases} \quad (4.2)$$

où les demi grand axe et demi petit axe A_{xk} et A_{yk} sont définis par :

$$\begin{cases} A_{xk} = \frac{P_k}{4\Sigma(e_k, \pi)} \\ A_{yk} = \frac{P_k \sqrt{1-e_k^2}}{4\Sigma(e_k, \pi)} \end{cases} \quad (4.3)$$

Pour déterminer $(x_g(t), y_g(t))$ en fonction de $(x_k(t), y_k(t))$, il est nécessaire de déterminer la fonction $\phi(t)$ pour laquelle la vitesse tangentielle $v_g(t)$ de l'ellipse géométrique est égale à la vitesse tangentielle $v_k(t)$ de l'ellipse cinématique, i.e. voir la Figure 4.1. En pratique, $v_g(t) = v_k(t)$ se met sous la forme d'une équation différentielle qui est résolue avec la méthode de Runge-Kutta d'ordre 4 et permet de résoudre un système d'équation non linéaire avec les conditions aux limites suivantes :

$$\begin{cases} \phi(0) = 0 \\ \phi(T) = 2\pi \end{cases} \quad (4.4)$$

avec T, P_g, e_g, e_{vk} et θ des paramètres fixés.

En résumé, un mouvement visuel est entièrement défini par un couple (e_g, e_{vk}) caractérisant la géométrie et la dynamique de la trajectoire respectivement. Dans les deux expériences, les stimuli visuels contiennent 19 périodes complètes de durée $T = 1.8$ s, impliquant que chaque stimulus dure 34.2 s. Six conditions visuelles ont ainsi été générées (2 formes géométriques) \times (3 cinématiques), voir Figure 4.2. Deux conditions respectent les mouvements biologiques, i.e. la loi en puissance 1/3 :

TABLE 4.1 – Minima et maxima des profils de vitesse des stimuli visuels (en cm.s^{-1}).

e_g	0 - Cercle			.9 - Ellipse		
e_{vk}	0	.9	-.9	0	.9	-.9
v_{min}	22.24	13.81	13.81	22.24	13.81	13.81
v_{max}	22.24	31.62	31.62	22.24	31.62	31.62

1. $(e_g = 0, e_{vk} = 0)$, qui correspond à un cercle parcouru à vitesse constante ;
2. $(e_g = .9, e_{vk} = .9)$, qui correspond à une ellipse parcourue en accélérant dans les parties plates et en ralentissant dans les parties courbes.

Les quatre autres conditions visuelles ne respectent pas les mouvements biologiques :

1. $(e_g = 0, e_{vk} = .9)$, le point lumineux accélère dans les parties horizontales du cercle ;
2. $(e_g = 0, e_{vk} = -.9)$, le point lumineux accélère dans les parties verticales du cercle ;
3. $(e_g = .9, e_{vk} = 0)$, la vitesse est constante le long de l'ellipse et des accélérations sont perçues dans les parties les plus courbées (Viviani et Stucchi, 1992) ;
4. $(e_g = .9, e_{vk} = -.9)$, la vitesse augmente dans les parties les plus courbées de l'ellipse et ralentit dans les parties les plus plates.

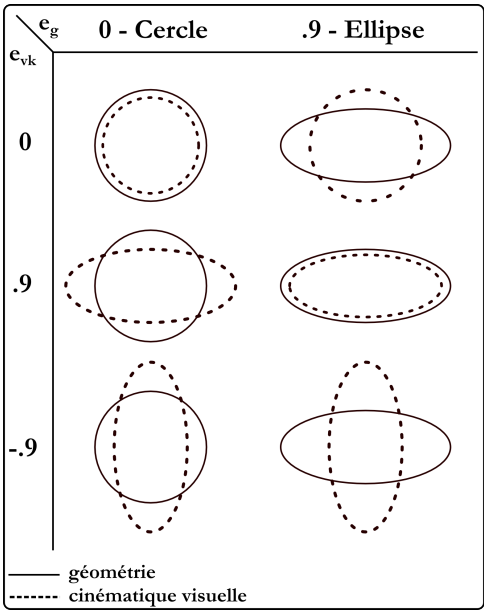


FIGURE 4.2 – Stimuli visuels : Les traits pleins représentent la géométrie de la trajectoire – les traits pointillés représentent la cinématique.

Les extrema des profils de vitesse sont présentés dans la table 4.1.

Stimuli audiovisuels.

🔊 Les stimuli audiovisuels ont été générés en combinant les 6 stimuli visuels de l'expérience 9 (cf. figure 4.2) avec 3 sons de frottement de synthèse évoquant le son d'une personne dessinant une

ellipse ou un cercle sur une feuille de papier. Afin de générer les sons, on utilise le même modèle de synthèse physiquement informé que dans les chapitres précédents (Gaver, 1993b; Van Den Doel *et al.*, 2001). Ce modèle qui simule une source sonore résultant d'une succession de micro-impacts d'un excitateur (ici un stylo) sur une surface rugueuse est brièvement rappelé ici. Le son est généré à partir d'un profil de vitesse correspondant à la vitesse tangentielle v_t du mouvement du stylo. En pratique la surface est modélisée par un bruit représentant la hauteur des aspérités de la surface. Ce bruit est ensuite filtré passe-bande avec une fréquence centrale qui varie proportionnellement à la vitesse tangentielle du stylo. Le filtre passe-bande est de type Butterworth avec un facteur de qualité constant de 3 (choisi arbitrairement). Dans le but de rajouter du réalisme, le bruit filtré est convolué par la réponse impulsionnelle d'un objet résonnant (Aramaki *et al.*, 2010). Le mapping entre la fréquence de coupure du filtre et la vitesse tangentielle du filtre est défini par $f_c(t) = \alpha v_t(t)$ où α est un coefficient de proportionnalité constant. Du point de vue perceptif, plus α est grand, plus le son généré sera brillant. Afin de choisir le mapping le plus approprié, son influence sur le geste produit dans une tâche de synchronisation audio-motrice a été évaluée dans une expérience préliminaire décrite en section 4.3¹. Le mapping a donc été fixé arbitrairement à $\alpha = 20$.

De la même façon que pour l'expérience visuelle, un profil de vitesse cinématique *sonore* (noté AK dans la suite) est défini pour chaque son et est ensuite utilisé pour contrôler le filtre du modèle de synthèse. Il correspond au profil de vitesse d'une personne parcourant un cercle ou une ellipse en respectant les mouvements biologiques. Par convention on notera e_{ak} l'excentricité de l'ellipse associée à ce profil de vitesse sonore. Il correspond donc au profil de vitesse des situations ($e_g = 0, e_{vk} = 0$) et ($e_g = .9, e_{vk} = .9$) définies dans le paragraphe précédent et qui ici seront notées $e_{ak} = 0$ et $e_{ak} = .9$. En plus de ces deux situations, une troisième situation $e_{ak} = -.9$ est définie relativement à l'orientation de l'ellipse visuelle. Finalement, 18 stimuli audiovisuels ont été générés en combinant les 6 stimuli visuels de l'expérience visuelle et les 3 sons de synthèse que l'on vient de définir. La figure 4.3 présente les 18 stimuli audio-visuels. Une condition, $e_g = e_{vk} = e_{ak}$, respecte les mouvements biologiques, toutes les autres conditions donnent lieu à des incongruences entre les mouvements visuels et ceux évoqués par les sons, en particulier :

1. $e_g = e_{vk} \neq e_{ak}$: les mouvements visuels respectent les mouvements biologiques et les informations de vitesse transmises par les sons sont incongruentes ;
2. $e_g = e_{ak} \neq e_{vk}$: les mouvements visuels ne respectent pas les mouvements biologiques alors que le mouvement évoqué par le son est congruent avec la géométrie de la forme visuelle ;
3. $e_g \neq (e_{vk} = e_{ak})$: les cinématiques évoquées par le mouvement visuel et par les sons sont congruentes mais les mouvements visuels ne respectent pas les mouvements biologiques ;
4. $e_g \neq e_{vk} \neq e_{ak}$: des accélérations sont perçues dans le mouvement visuel alors que le son évoque soit des décélérations (soit une vitesse constante), et vice-versa.

4.2.3 Appareillage

Les deux expériences se sont déroulées dans le noir dans une cabine audiométrique. Les sujets étaient assis devant un écran d'ordinateur DELL 1907fp dont la résolution était de 1280 x 1024 pixels et devant une tablette graphique Wacom Intuos5. Afin de ne pas voir leurs mains, un cache

1. Des expériences plus détaillées présentées en annexe 7.2 montrent que le mapping n'a pas d'influence sur la taille de l'ellipse ou du cercle dessiné.

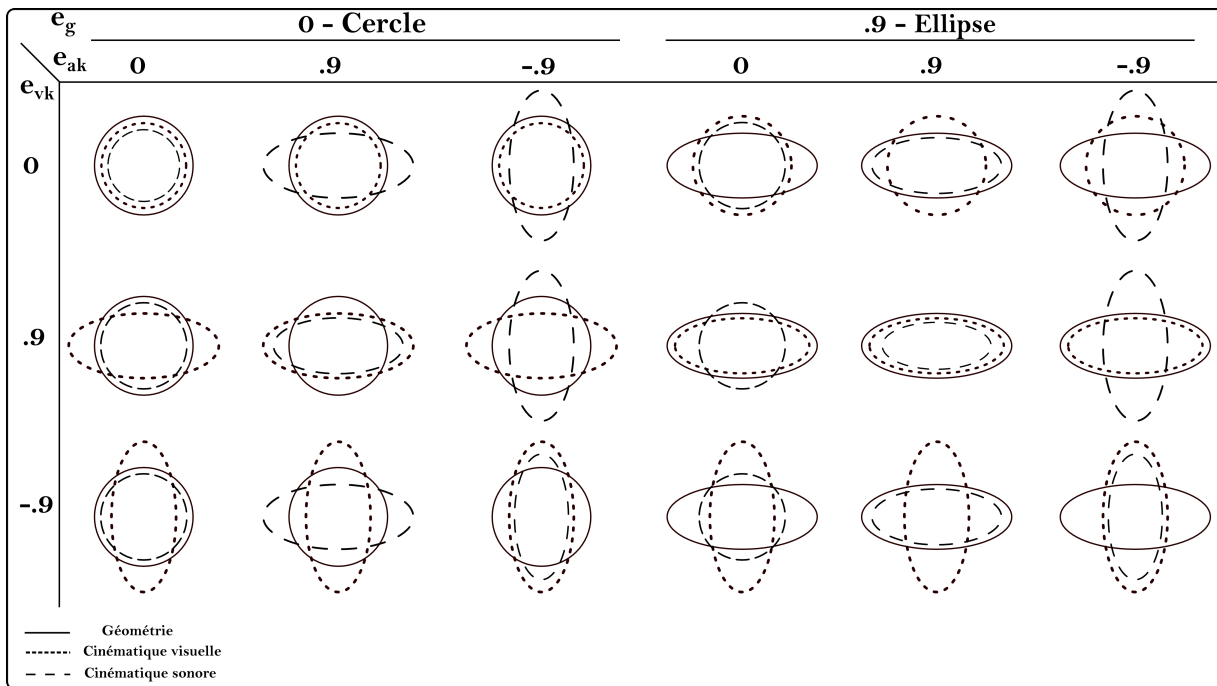


FIGURE 4.3 – Stimuli audio-visuels : Les traits pleins représentent la géométrie de la trajectoire – les petits traits pointillés représentent la cinématique visuelle – les grands très pointillés représentent la cinématique sonore.

était disposé au-dessus de la tablette graphique (voir figure 4.4(a)). Dans l'expérience audiovisuelle, les sujets portaient également un casque Sennheiser-HD650 pour la diffusion des sons (voir figure 4.4(b)). Le volume sonore était ajusté afin que le niveau d'écoute soit confortable. L'expérimentateur était présent dans la cabine durant toute la durée de l'expérience afin de contrôler le déclenchement des stimuli visuels et sonores. Les stimuli étaient générés en temps-réel sur un ordinateur iMac par le biais du logiciel Max/MSP. Les stimuli visuels étaient affichés sur l'écran DELL via une interface DVI, le taux de rafraîchissement de l'écran était de 60 Hz et le point lumineux blanc était généré en affichant séquentiellement une série de 108 points afin de percevoir un mouvement le plus fluide possible. Les sons étaient présentés avec un taux d'échantillonnage de 44100 Hz. Les mouvements graphiques étaient eux enregistrés par une tablette graphique Wacom Intuos5 avec un taux d'échantillonnage de 129 Hz et une précision spatiale de 5.10^{-3} mm.

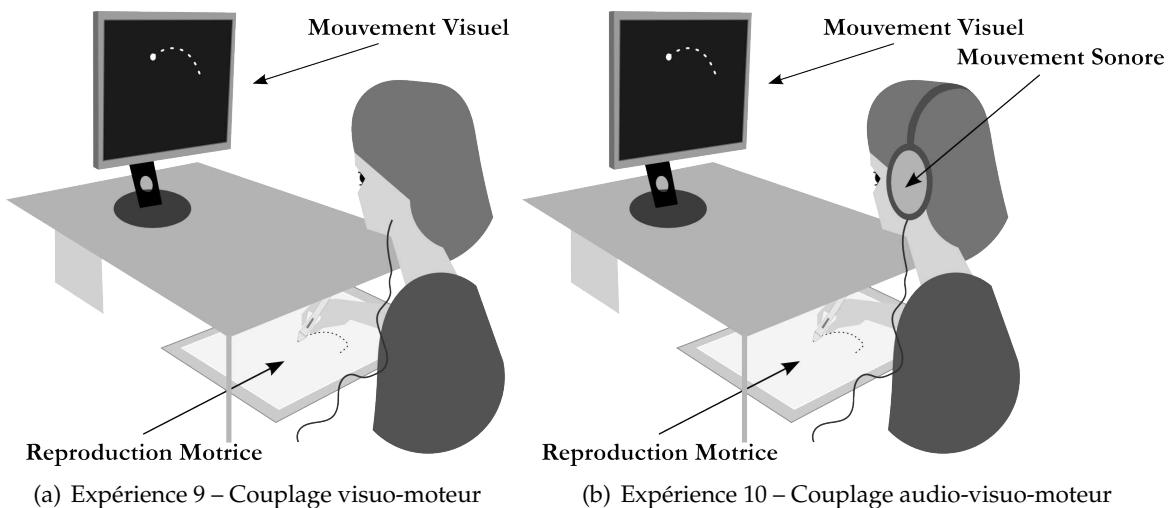


FIGURE 4.4 – Dispositifs expérimentaux des deux expériences

4.2.4 Tâche

Les deux expériences ont été réalisées dans le même ordre pour tous les sujets, d’abord l’expérience visuelle et ensuite l’expérience audiovisuelle. Chaque expérience était précédée d’une phase de familiarisation durant laquelle les sujets étaient informés qu’ils allaient devoir synchroniser leurs gestes à l’aide d’un stylet sur une tablette graphique avec les mouvements visuels utilisés dans l’expérience 9 (respectivement audiovisuels dans l’expérience 10). La tâche de familiarisation durait le temps nécessaire pour que les sujets aient bien compris la consigne, ce qui correspondait généralement à deux ou trois essais. Il était demandé aux sujets de synchroniser leurs gestes dans le sens inverse des aiguilles d’une montre pour avoir l’impression qu’ils étaient en train de contrôler eux-même le mouvement du point du lumineux (et du son de frottement dans le cas de l’expérience audiovisuelle). Pour générer le mouvement, il leur était demandé de bloquer leur poignet et de lever leur coude afin d’utiliser uniquement l’articulation épaule-coude dans la génération du mouvement. Enfin, il leur était demandé de bloquer leur torse et plus précisément leur tronc afin d’éviter que des oscillations posturales influencent le mouvement produit.

4.2.5 Analyse des données

Dans les deux expériences, les enregistrements des productions motrices ont été analysés selon leurs caractéristiques géométriques et temporelles. Les données enregistrées correspondent aux coordonnées du stylet sur la tablette graphique échantillonnées au cours du temps (x_n, y_n) . 19 cycles complets de mouvements ont été enregistrés à chaque essai. A chaque essai, les sujets avaient besoin d’un certain nombre de cycles pour se synchroniser avec le stimulus visuel ou audio-visuel. Seules 10 périodes, entre la sixième et la seizième période (inclue), ont ainsi été utilisées pour l’analyse des données. Les données brutes ont été lissées afin d’éliminer le bruit numérique lié au taux d’échantillonnage élevé en appliquant un filtre de Savitsky-Golay (Savitzky et Golay, 1964) avec une fenêtre de 43 échantillons et une interpolation du troisième ordre, ce qui est équivalent à un filtrage passe-bas avec une fréquence de coupure de 8 Hz. De plus, un filtrage passe-haut (Butterworth) avec une fréquence de coupure de 0.2 Hz a été appliqué afin d’éliminer la dérive spatiale des sujets au cours du mouvement. Le centre de gravité des formes dessinées se déplaçait en effet au cours du temps parce que les sujets ne voyaient pas ce qu’ils faisaient.

Afin de déterminer les caractéristiques géométriques de la forme perçue et reproduite, les performances motrices des sujets ont été analysées de la façon suivante. Pour chaque période de chaque essai, l’ellipse qui représente le mieux la trace dessinée a été déterminée en utilisant une méthode classique considérant que le nuage de points (x_n, y_n) est un ensemble de masses de même poids (Viviani *et al.*, 1997; Goldstein, 1962). En diagonalisant le tenseur d’inertie de ce nuage de points, il est possible de déterminer les caractéristiques géométriques de cette ellipse, i.e. son excentricité, sa taille et son inclinaison. Les valeurs propres sont proportionnelles aux demi grand axe et demi petit axe, les vecteurs propres permettent de déterminer l’inclinaison de l’ellipse. Cette méthode est un moyen efficace pour déterminer l’ellipse qui s’ajuste le mieux au nuage de points au sens des moindres carrés. A partir des demi grands axes et demi petits axes, l’excentricité de l’ellipse a ainsi pu être calculée et moyennée à travers 10 valeurs pour chaque essai.

Les aspects temporels des performances motrices présentent également un intérêt pour remonter au mouvement qui a été perçu. En effet, si notre intérêt principal porte sur la géométrie de la forme qui a été perçue et reproduite, la cinématique du mouvement réalisé peut également nous aider à comprendre d’éventuelles distorsions. Comme on l’a vu en introduction de ce chapitre et dans l’état de l’art, la géométrie d’un mouvement graphique est intimement liée à sa dynamique. Nous nous

sommes donc focalisés sur les différences entre le profil de vitesse du geste produit et le profil de vitesse de la cible visuelle dynamique. Le but de cette comparaison n'est pas d'évaluer la synchronie entre les profils, les sujets pouvant avoir soit de l'avance soit du retard par rapport à la cible visuelle, mais de déterminer si le profil a été bien reproduit dans *sa forme*². Par ailleurs, les performances motrices des sujets diffèrent en taille et en inclinaison. Il est donc nécessaire de normaliser les profils de vitesse en amplitude avant d'effectuer des comparaisons. Afin d'effectuer une normalisation cohérente entre les sujets, chaque enregistrement (x_n, y_n) a été normalisé par rapport au périmètre de la forme cible. Enfin, une rotation a été effectuée pour obtenir une ellipse horizontale dont le demi-grand axe est aligné sur l'axe des abscisses pour pouvoir calculer les valeurs du demi-grand et demi-petit axe. Afin de quantifier la différence entre les profils de vitesse des sujets et les profils de vitesse cibles, une distance basée sur une comparaison des spectres des profils a été réalisée. En effet, comparer les spectres permet de s'affranchir d'un éventuel délai temporel positif ou négatif entre les deux profils. La différence entre les spectres est alors calculée en utilisant une distance euclidienne qu'on appellera par la suite distorsion cinématique (KD) :


$$KD = \sqrt{\frac{\sum_i (|\hat{v}_{ref}|_i - |\hat{v}_{exp}|_i)^2}{\sum_j |\hat{v}_{ref}|_j^2}} \quad (4.5)$$

où \hat{v}_{ref} et \hat{v}_{exp} correspondent respectivement aux spectres du profil de vitesse de la cible visuelle et du sujet. Afin de réaliser des comparaisons entre les profils de vitesse des différentes formes, cette distance est normalisée par l'énergie du profil de vitesse de la cible $\sum_j |\hat{v}_{ref}|_j^2$.

4.2.6 Statistiques

Dans les deux expériences, des analyses statistiques séparées ont été réalisées en fonction de la géométrie de la trajectoire parcourue (e_g), circulaire ou elliptique. Dans l'expérience 9 (visuelle), le plan d'analyse statistique comporte un facteur expérimental, la cinématique visuelle (VK) avec trois modalités. Dans l'expérience 10 (audiovisuelle), le plan d'analyse statistique a deux niveaux : 3 Cinématiques visuelles (VK) x 3 Cinématiques Sonores (AK). Afin d'évaluer les effets de chaque facteur expérimental sur l'excentricité et la distorsion cinématique, des ANOVA à mesures répétées ont été réalisées. Préalablement à ces analyses, nous avons vérifié les pré-requis statistiques de normalité et d'homogénéité des variances de nos échantillons. Chaque effet significatif a été analysé en détail en utilisant un test post-hoc de Newman-Keuls. Le niveau de significativité a été fixé à .05 pour toutes les analyses.

4.3 Expérience préliminaire – Couplage audio-moteur

 Le but de cette expérience était de déterminer le mapping le plus approprié pour synthétiser les sons de frottement utilisés dans l'expérience audiovisuelle. Cette expérience a été réalisée sur les mêmes sujets que pour les deux autres expériences, et a été réalisée systématiquement avant les deux autres. Afin d'évaluer dans quelle mesure le mapping entre la vitesse tangentielle du mouvement et la fréquence de centrale du filtre passe-bande peut influencer la performance motrice, on a demandé aux sujets de synchroniser leurs mouvements avec des sons de synthèse générés à partir de différents mappings avec les yeux bandés (i.e. en boucle ouverte visuelle). Les stimuli étaient composés de sons de synthèse correspondant soit au cercle soit à l'ellipse définis précédemment (i.e. $e_{ak} = 0$ ou $e_{ak} = .9$) et à partir de 3 mappings ($\alpha_1 = 10$, $\alpha_2 = 20$, $\alpha_3 = 40$), les amplitudes de variation des fréquences

2. Par reproduire la forme du profil de vitesse on entend le fait d'avoir été capable de reproduire le profil de vitesse du stimulus cible avec une avance ou un retard éventuel.

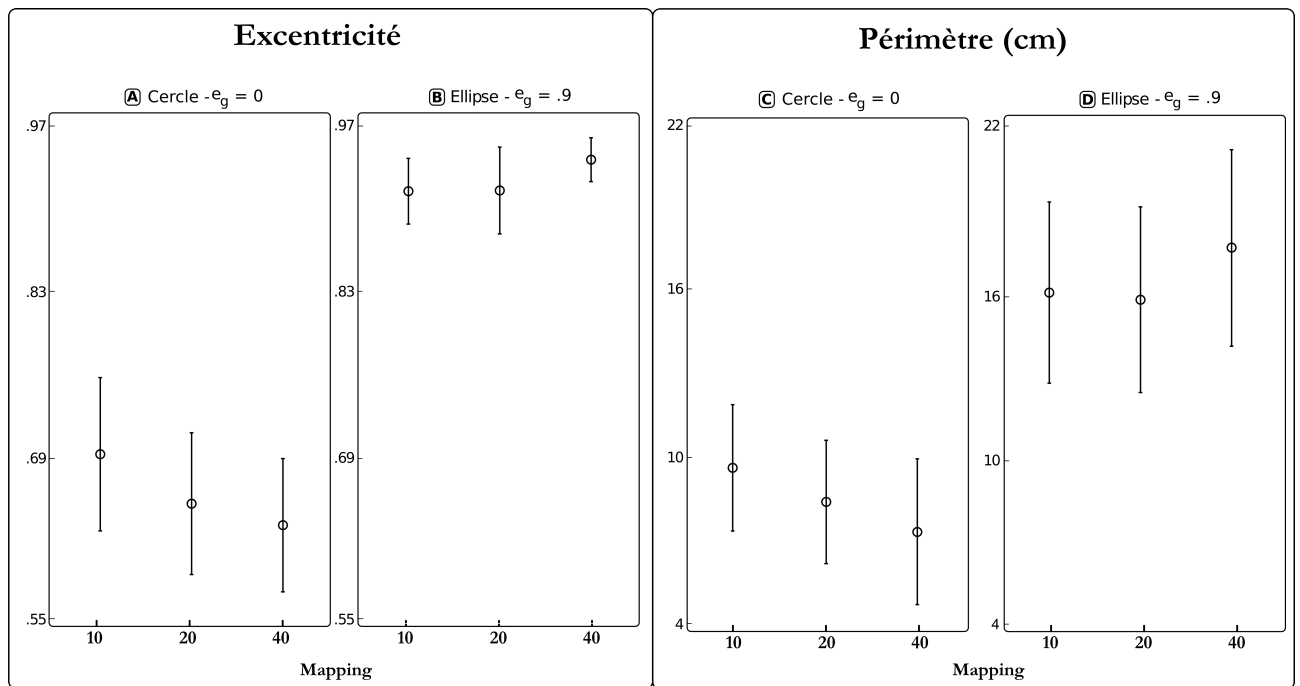


FIGURE 4.5 – Résultats de l'expérience de synchronisation audio-motrice : excentricités et périmètres (en cm) – les barres d'erreurs représentent l'erreur standard.

de coupure sont présentées dans la table 4.2. Les résultats n'ont pas mis en évidence d'influence

TABLE 4.2 – Fréquences centrales minimums et maximums (en Hz) des profils de vitesse pour chacun des mappings – dans le cas du cercle la fréquence centrale est constante durant toute la durée du mouvement

e_{ak}	0 - Cercle	.9 et -.9 - Ellipse	
Mapping (α)	f	f_{min}	f_{max}
10	222.4	138.1	316.2
20	444.9	276.2	632.3
40	889.8	552.4	1264.6


significative du mapping sur l'excentricité des formes reproduites ni sur le périmètre des formes dessinées par les sujets. (Excentricités – (Cercle) : $F(2,32) = 2.350$, $p = .11$ – (Ellipse) : $F(2,32) = 1.834$, $p = .1876$; Périmètres : (Cercle) : $F(2,32) = 3.252$, $p = .0651$; (Ellipse) : $F(2,32) = 3.009$, $p = .07634$; voir Figure 4.5). En se basant sur ces résultats, le mapping $\alpha = 20$ a été choisi car c'est celui qui procure le plus grand réalisme par rapport à un vrai son de frottement (suite à des discussions avec les sujets). Une étude spécifique sur l'influence du mapping et sa relation avec les caractéristiques géométriques des performances motrices des sujets dans une même tâche de synchronisation est présentée en annexe 7.2.

4.4 Expérience 9 – Couplage visuo-moteur

Dans cette première expérience, le but est d'évaluer si le couplage visuo-moteur en boucle ouverte entre un mouvement graphique et une stimulation visuelle est bien contraint par la covariation entre la vitesse du point lumineux et la courbure de la trajectoire parcourue. Les sujets devaient donc synchroniser leurs gestes avec des stimuli visuels sans voir leurs mains (cf. Figure 4.4(a)). Les sti-

multi visuels étaient composés de mouvements circulaires ou elliptiques respectant ou non la loi en puissance 1/3 (cf. Figure 4.2). Au regard des connaissances sur la perception et reproduction des mouvements biologiques, on souhaite ici vérifier l'influence de la dynamique du mouvement visuel sur la géométrie de la trajectoire perçue et reproduite et plus exactement si la cinématique visuelle peut distordre la forme perçue quand elle ne respecte pas un mouvement biologique. Cela a été mis en évidence au niveau de la perception visuelle (Viviani et Stucchi, 1989), et de la perception kinesthésique (Viviani *et al.*, 1997). Viviani *et al.* (1987) ont également montré ces résultats dans une tâche de couplage kinesthésico-moteur en boucle fermée. Ici on s'intéresse à cette situation dans le cas d'un tracking visuo-moteur en boucle ouverte, situation qui n'a jusqu'alors pas été étudiée à notre connaissance.

4.4.1 Résultats

Les résultats de cette expérience sont présentés par descripteur (excentricité et distorsion cinématique) et sont ensuite discutés. La table 4.3 et la figure 4.6 présentent l'ensemble des moyennes et erreurs standards de chaque descripteur par facteur. Une vidéo présentant des exemples de reproduction motrice typiques est disponible sur la page web associée à la thèse. 

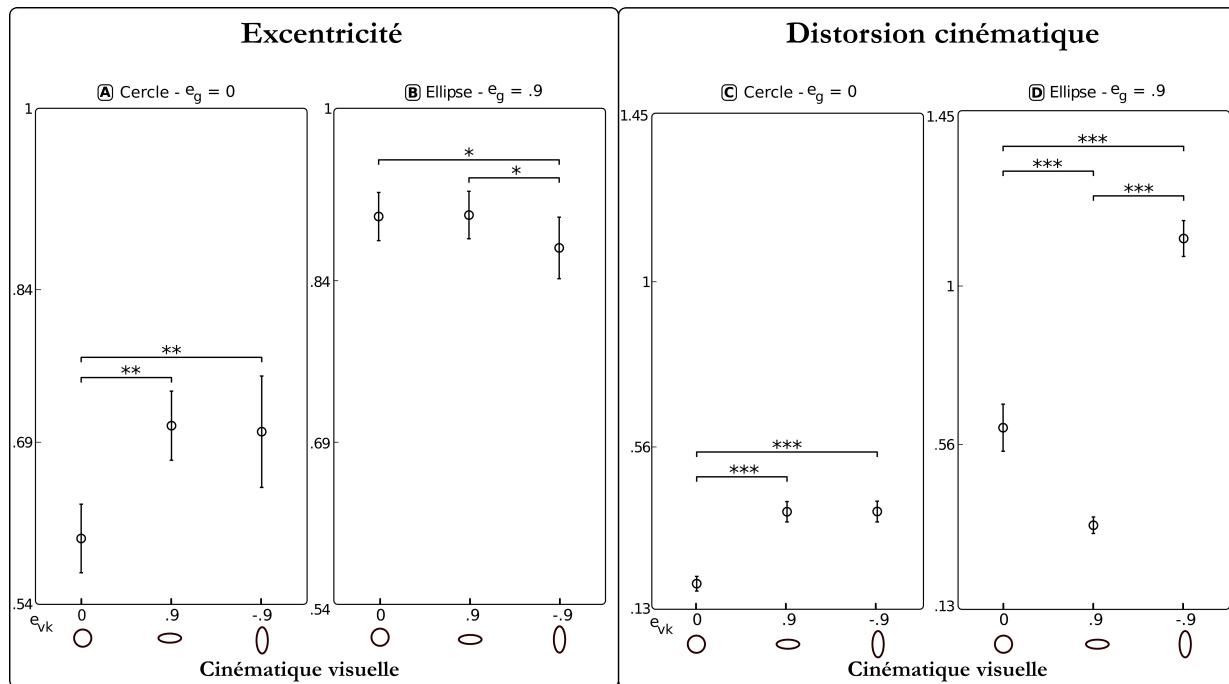


FIGURE 4.6 – Résultats de l'expérience de synchronisation visuo-motrice : excentricités et périmètres (en cm) – les barres d'erreurs représentent l'erreur standard — *** : $p < .001$; ** $p < .01$; * $p < .05$.

Excentricité.

Les résultats mettent en évidence que la cinématique visuelle modifie significativement l'excentricité des reproductions motrices à la fois pour les cercles et pour les ellipses ($F(2, 32) = 9.40, p < .001$ et $F(2, 32) = 3.94, p = .02$ respectivement) :

- **Pour les cercles**, dans les situations non biologiques, i.e. cercle parcouru avec une cinématique elliptique, $e_{vk} = .9$ et $e_{vk} = -.9$, ils sont reproduits avec une excentricité significativement plus grande que dans la condition biologique $e_{vk} = 0$ (Newman-Keuls : $p < .01$ et $p < .01$ respectivement).
- **Pour les ellipses**, dans la situation non-biologique $e_{vk} = -.9$, i.e. déphasage de 45 degrés entre

TABLE 4.3 – Moyennes et erreurs standards de l'excentricité et de la distorsion cinématique pour chaque forme et chaque condition visuelle

e_g	0 - Cercle			.9 - Ellipse		
e_{vk}	0	.9	-.9	0	.9	-.9
M_{Ecc}	.60	.70	.70	.90	.90	.87
SE_{Ecc}	.016	.016	.026	.011	.011	.014
M_{KD}	.19	.39	.39	.60	.34	1.11
SE_{KD}	.010	.014	.014	.032	.011	.024

l'ellipse géométrique et l'ellipse cinématique, les ellipses sont reproduites de façon significativement plus arrondies que dans les situations $e_{vk} = 0$ et $e_{vk} = .9$, i.e. forme elliptique parcourue avec une cinématique circulaire ou elliptique déphasée (Newman-Keuls : $p < .05$ et $p < .05$ respectivement). On peut noter que seule la situation non biologique $e_{vk} = -.9$ a été reproduite significativement plus arrondie que la situation $e_{vk} = .9$, $e_{vk} = 0$ et n'a pas entraîné de modification significative de l'excentricité.

Distorsion cinématique.

Les analyses de la distorsion cinématique viennent appuyer les résultats précédents sur l'excentricité, en effet, il est clair que la cinématique visuelle a modifié la géométrie des reproductions motrices. La distorsion cinématique nous renseigne elle sur la capacité que les sujets ont eu à reproduire la cinématique du mouvement, indépendamment de la forme et d'un retard par rapport à la cible.

Les résultats mettent en évidence que la cinématique visuelle a significativement affecté cette capacité à reproduire le mouvement pour les deux cercles et les ellipses ($F(2, 32) = 67.42$, $p < .001$ et $F(2, 32) = 204.70$, $p < .001$ respectivement) et qu'en particulier la distorsion cinématique a été significativement plus grande aussi bien pour les cercles que pour les ellipses dans les situations non-biologiques et dans les situations biologiques, (Newman-Keuls : cercles : $p < .001$ pour $e_{vk} = .9$ et $p < .001$ pour $e_{vk} = -.9$; ellipse : $p < .001$ pour $e_{vk} = 0$ et $p < .001$ pour $e_{vk} = -.9$). Cela peut donc expliquer pourquoi la géométrie des formes reproduites a été significativement modifiée dans les situations non-biologiques (sauf pour $e_{vk} = 0$ dans le cas des ellipses), les sujets ont pu essayer d'intégrer à la fois la forme et le mouvement dans leur reproduction, et comme ils ne voyaient pas le geste qu'ils étaient en train de faire, la cinématique a contraint la géométrie de la forme dessinée.

4.4.2 Discussion

Ces résultats viennent donc appuyer notre hypothèse sur l'influence de la cinématique visuelle sur la perception et la reproduction des mouvements biologiques, ils confirment également les différentes études de Viviani sur la perception visuelle (Viviani et Stucchi, 1989) et reproduction en boucle visuelle fermée (Viviani et Mounoud, 1990; Viviani *et al.*, 1987) d'un mouvement.

Cela est particulièrement flagrant dans le cas des trajectoires circulaires parcourues avec une cinématique non-biologique correspondant à une trajectoire elliptique ($e_{vk} = .9$ et $e_{vk} = -.9$), les sujets étant obligés d'accélérer bien que la courbure de la trajectoire reste constante. Leur geste semble donc *entraîné* à parcourir une trajectoire plus plate que celle qu'ils pensent réaliser.

Dans le cas des ellipses, les excentricités ne sont significativement différentes qu'avec la situation $e_{vk} = -.9$, i.e. situation où la cinématique visuelle est inversée par rapport à la situation biologique,

la vitesse tangentielle du point lumineux augmente au lieu de diminuer dans les parties les plus courbes de la trajectoire et réciproquement. Cependant, l'analyse de la distorsion cinématique met clairement en évidence que, bien que la géométrie de la trajectoire ait été *respectée*, les sujets ont quand même eu beaucoup plus de mal à suivre le mouvement du point dans les situations non-biologiques.


Ces études viennent également compléter les études de Paolo Viviani et ses collègues (Viviani et Stucchi, 1989, 1992; Viviani *et al.*, 1997). En effet, bien qu'un tel effet d'une cinématique visuelle sur la perception de la géométrie d'une trajectoire ait déjà été mis en évidence dans de nombreuses expériences, la situation de reproduction en boucle ouverte visuelle n'avait jusqu'alors pas été étudiée. Cela confirme donc que l'intégration de la géométrie d'un mouvement est contrainte par la relation entre la vitesse du mouvement et la courbure de la trajectoire, et que, pour que la géométrie de la trajectoire soit intégrée de façon optimale, les covariations vitesse/courbure doivent respecter la loi en puissance 1/3.

4.5 Expérience 10 – Couplage audio-visuo-moteur

Les résultats de l'expérience précédente ont donc confirmé que la perception et la reproduction d'une trajectoire visuelle étaient contraintes par les covariations entre cinématique et courbure de la trajectoire. L'objectif initial de cette étude était d'évaluer l'influence de sons évoquant des mouvements sur la perception et reproduction de mouvements visuels. La même expérience que l'expérience visuelle a donc été reproduite mais avec des stimuli audiovisuels combinant les stimuli visuels de l'expérience 9 et des sons de frottement de synthèse évoquant les mouvements produits par une personne qui dessine différentes trajectoires. En associant des mouvements circulaires ou elliptiques dans les deux modalités, 18 combinaisons, congruentes ou incongruentes, entre des stimuli visuels et auditifs ont ainsi été générées (voir figure 4.3).

L'objectif de cette expérience était d'évaluer dans quelle mesure l'ajout d'un son évoquant une trajectoire visuelle peut modifier la géométrie de la forme visuelle perçue et de sa reproduction motrice. En effet, l'expérience précédente a mis en évidence que la dynamique visuelle d'un mouvement biaisait la perception et la reproduction la forme perçue. L'hypothèse évaluée ici est la suivante : si la modalité auditive a la capacité d'évoquer très fidèlement une dynamique, et de façon plus précise que la modalité visuelle, combiner une dynamique sonore à une dynamique visuelle (biologique ou non) devrait alors soit amplifier la distorsion visuelle si le mouvement visuel est déjà non biologique, soit donner lieu à une nouvelle illusion audiovisuelle quand le mouvement visuel est biologique. Dans ce cas, le son devrait prendre le dessus sur la modalité visuelle en modifiant la géométrie de la forme perçue et reproduite.

4.5.1 Résultats

Les résultats de cette expérience sont présentés par descripteur (excentricité et distorsion cinématique) et sont ensuite discutés. La table 4.4 et la figure 4.7 présentent l'ensemble des moyennes et erreurs standards de chaque descripteur par facteur. Une vidéo présentant des exemples de reproduction motrice typiques est disponible sur la page web associée à la thèse. 

Excentricité.

Les résultats mettent en évidence l'influence du son sur la perception visuelle et la reproduction motrice. Un effet principal de la cinématique sonore sur l'excentricité des formes reproduites, que ce soit pour les cercles ($F(2, 32) = 9.26, p < .001$) que pour les ellipses ($F(2, 32) = 4.35, p < .05$), a été mis en évidence :

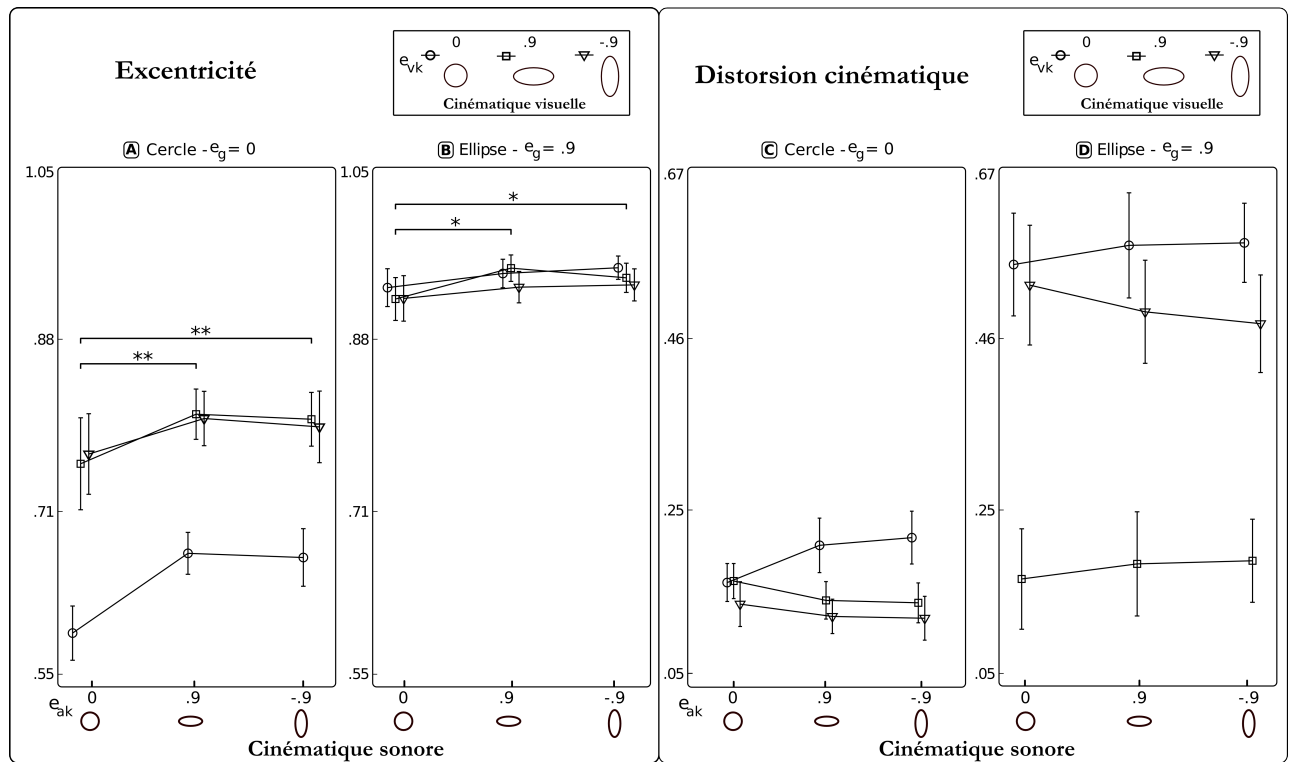


FIGURE 4.7 – Résultats de l'expérience de synchronisation audio-visuo-motrice : excentricités et périmètres (en cm) – les barres d'erreurs représentent l'erreur standard — *** $p < .001$; ** $p < .01$; * $p < .05$, les significativités correspondent à l'effet principal de la cinématique sonore.

TABLE 4.4 – Moyennes et erreurs standards de l'excentricité et de la distorsion cinématique pour chaque forme et chaque condition audiovisuelle

e_g	0 - Cercle								
e_{vk}	0			.9			-.9		
e_{ak}	0	.9	-.9	0	.9	-.9	0	.9	-.9
M_{Ecc}	.59	.67	.66	.75	.81	.80	.77	.80	.79
SE_{Ecc}	.013	.010	.014	.023	.012	.013	.020	.014	.018
M_{KD}	.16	.20	.21	.16	.14	.13	.13	.12	.12
SE_{KD}	.012	.017	.016	.010	.012	.012	.014	.010	.014
e_g	.9 - Ellipse								
e_{vk}	0			.9			-.9		
e_{ak}	0	.9	-.9	0	.9	-.9	0	.9	-.9
M_{Ecc}	.93	.94	.95	.92	.95	.94	.92	.93	.93
SE_{Ecc}	.009	.007	.006	.010	.007	.007	.011	.008	.008
M_{KD}	.55	.57	.58	.16	.18	.19	.52	.49	.47
SE_{KD}	.032	.033	.025	.031	.032	.026	.037	.032	.030

- Les cercles ont été reproduits de façon significativement plus plate, indépendamment de la cinématique visuelle, dans les situations où le son évoquait des accélérations dans deux parties opposées du cercle, i.e. $e_{ak} = .9$ et $e_{ak} = -.9$ (Newman-Keuls : $p < .05$) ;
- Les ellipses ont été reproduites de façon plus aplatie, indépendamment de la cinématique visuelle, dans les situations $e_{ak} = .9$ (i.e. biologique) et $e_{ak} = -.9$ (i.e. non-biologique) que dans la situation $e_{ak} = 0$ (i.e. non biologique) (Newman-Keuls : $p < .05$).

Ces résultats viennent donc appuyer notre hypothèse initiale sur le fait que les variations de timbre évoquées des stimuli sonores modifient clairement la géométrie reproduite dans les situations audiovisuelles.

Les résultats montrent également un effet principal de la cinématique visuelle sur l'excentricité de l'ellipse reproduite. L'excentricité des reproductions motrices est en effet significativement affectée par la condition cinématique visuelle quelle que soit la cinématique sonore aussi bien pour les cercles ($F(2, 32) = 111.45, p < .001$) que pour les ellipses ($F(2, 32) = 5.86, p < .01$).

- **Les cercles** dessinés ont été significativement aplatis quand la cinématique visuelle ne respectait pas les mouvements biologiques (Newman-Keuls : $p < .001$ pour $e_{vk} = .9$; $p < .001$ pour $e_{vk} = -.9$).
- **Les ellipses** ont elles été reproduites de façon plus arrondie dans la situation non-biologique $e_{vk} = -.9$ que dans les situations non-biologique circulaire $e_{vk} = 0$, ou biologique $e_{vk} = .9$ (Newman-Keuls : $p < .01$ et $p < .05$ respectivement).

Ces résultats sont cohérents avec ceux de l'expérience visuelle précédente. Cependant, il est intéressant de constater que les effets sont encore plus forts dans cette expérience que dans l'expérience visuelle. Dans l'expérience visuelle, les cercles ont été reproduits dans les situations non-biologiques, i.e. $e_{vk} = .9$ et $e_{vk} = -.9$, avec une excentricité moyenne de .7, tandis que dans l'expérience audiovisuelle, les conditions non-biologiques ont été reproduites avec une excentricité en moyenne égale à .8. Cela montre donc que le son a amplifié l'illusion de la situation purement visuelle.

Finalement, l'interaction entre cinématique visuelle et cinématique sonore est également significative uniquement dans le cas des ellipses ($F(4, 64) = 2.81, p < .05$). Pour chaque cinématique visuelle, la condition $e_{ak} = 0$, i.e. vitesse constante, implique des ellipses significativement plus rondes que dans les deux autres conditions (Newman-Keuls : $p < .05$ pour toutes les comparaisons). Il est également important de noter que dans le cas de la cinématique visuelle biologique $e_{vk} = .9$, $e_{vk} = 0$ arrondit également l'ellipse par rapport aux deux autres conditions $e_{ak} = .9$ et $e_{ak} = -.9$ (Newman-Keuls : $p < .001$ pour les deux comparaisons).

Distorsion cinématique.

Comme pour l'expérience visuelle, l'analyse de la distorsion cinématique aide à mieux interpréter les résultats sur les excentricités. Cette distorsion nous renseigne sur la qualité de la reproduction de la cinématique du mouvement visuel sans tenir compte de l'avance ou du retard du geste par rapport au point lumineux.

On aurait pu s'attendre à ce que la distorsion cinématique soit plus ou moins importante en fonction de la condition sonore, cependant aucun effet de la cinématique sonore n'a été mis en évidence ($F(2, 32) < 1$ pour les deux formes).

Cependant, comme pour l'excentricité, la distorsion cinématique est significativement différente selon les conditions visuelles pour les deux formes (Cercle : $F(2,32) = 13.78, p < .001$; Ellipse : $F(2,32) = 42.81, p < .001$).

- **Dans le cas des ellipses**, la distorsion cinématique est moins grande pour la condition biologique $e_{vk} = .9$ que pour les deux autres conditions ($p < .001$ pour toutes les comparaisons).
- **A l'inverse, pour les cercles**, la distorsion cinématique est particulièrement modifiée dans le cas de cinématique biologique $e_{vk} = 0$, en particulier quand un son incongruent est présenté (i.e. $e_{ak} = .9$ et $e_{ak} = -.9, p < .05$, l'interaction entre cinématique visuelle et cinématique sonore est significative ($F(4,64) = 3.92, p < .001$)).

Finalement, bien que la cinématique sonore n'ait pas influencé de la même façon toutes les productions, l'analyse de l'interaction, en particulier dans le cas des cercles, montre que les sujets ont basé leurs reproductions principalement sur les informations transmises par la modalité sonore. En effet, dans les situations non-biologiques, la distorsion cinématique est plus faible en présence d'un son congruent évoquant des accélérations, alors que quand la cinématique respecte les mouvements biologiques, la distorsion cinématique est plus grande pour les situations non congruentes.

4.5.2 Discussion

Finalement, cette expérience a mis en évidence que le son affectait de façon impressionnante le couplage visuo-moteur avec un mouvement visuel. Cette influence est différente dans le cas des mouvements circulaires et des mouvements elliptiques. Les cercles sont clairement aplatis par des cinématiques sonores elliptiques quel que soit le type de cinématique visuelle. Et de plus, quand la cinématique visuelle est déjà biaisée, i.e. non-biologique, alors l'illusion visuelle est renforcée et les ellipses sont encore plus aplaties. Cela est confirmé par la diminution de la distorsion cinématique dans ces situations et montre l'importance de la corrélation entre les informations acoustiques et sonores dans les processus d'intégration multisensoriels (Parise *et al.*, 2012). Dans le cas des ellipses, la géométrie de la reproduction des formes a été significativement affectée quelle que soit la cinématique visuelle, et en particulier, l'ajout d'un son évoquant une vitesse constante a significativement arrondi les reproductions motrices. La distorsion cinématique montre bien les mêmes effets que dans l'expérience visuelle au niveau de l'influence des cinématiques visuelles non-biologiques sur la capacité à reproduire le mouvement. On aurait pu s'attendre à ce que les mêmes effets apparaissent pour les cinématiques sonores incongruentes, cependant il n'y a pas d'effet clair de la cinématique sonore sur la cinématique du mouvement dans ces situations.

Il est notable que l'illusion mise en évidence ici est bien plus saillante dans le cas des formes circulaires que des formes elliptiques. Ce résultat est à mettre en perspective avec les connaissances biomécaniques sur les mécanismes sous-jacents à de tels mouvements circulaires ou elliptiques. Il a été montré que les mouvements humains étaient modélisables par des systèmes dynamiques mettant en jeu des états stables, aussi appelés attracteurs moteurs, qui contraignent nos mouvements et en particulier les mouvements répétitifs. Sallagoity *et al.* (2004) et Athenes *et al.* (2004) ont en particulier mis en évidence que nous dessinons préférentiellement des formes elliptiques plutôt que des cercles. Cela pourrait expliquer pourquoi les effets du son ont été moindres dans le cas des ellipses que dans le cas des cercles. Ce point sera discuté dans le cadre général des deux expériences dans la discussion qui suit.

Par dessus tout, cette expérience montre de façon évidente que le son peut substantiellement affecter la modalité visuelle, et cela, dans une situation où on pourrait *a priori* penser que la vision

serait plus performante que l'audition (i.e. voir et reproduire une cible visuelle dynamique). Jusqu'à présent, les interactions entre vision et audition ont principalement été étudiées en utilisant des stimuli discrets spatialisés (Soto-Faraco *et al.*, 2002, 2003, 2004), et il est bien connu que la modalité visuelle est plus adaptée dans la localisation d'une source ponctuelle que la modalité auditive. Seules quelques études ont pu montrer qu'un stimulus discret pouvait interférer, et modifier la perception d'un stimulus visuel (Sekuler et Lau, 1997; Shams *et al.*, 2000) en utilisant la meilleure résolution temporelle de l'audition pour percevoir un stimulus discret. Dans le contexte sensori-moteur, Repp et Penel (2004) ont eux mis en évidence dans des tâches de tapping que les rythmes sonores attiraient plus les mouvements que les rythmes visuels. L'illusion mise en évidence ici utilise elle la capacité de l'audition à transmettre des informations dynamiques par le biais de ses variations intrinsèques de timbre ce qui nous oblige à reconsidérer le rôle de l'audition dans les processus d'intégration audiovisuels des mouvements continus, processus dans lesquels l'audition semble jouer jusqu'à présent un rôle mineur. Cela sera discuté dans la discussion générale qui suit.

4.6 Discussion générale

Dans l'étude présentée ici, la capacité à percevoir et reproduire un mouvement visuel, biologique ou non, a été étudiée en deux temps. Premièrement dans une expérience en boucle ouverte visuelle et dans un second temps dans une expérience en boucle ouverte audiovisuelle. Le but de la première expérience était d'évaluer dans quelle mesure le couplage visuo-moteur entre un mouvement visuel, biologique ou non, et sa reproduction en boucle ouverte était contraint par une relation entre vitesse et courbure, i.e. la loi en puissance $1/3$. Le second but était d'évaluer dans quelle mesure l'ajout d'un son de frottement évoquant une dynamique pouvait perturber, et même modifier, la perception et reproduction des mêmes mouvements, biologiques ou non. Pour cela, des combinaisons congruentes et non congruentes entre des sons continus et des mouvements visuels ont été réalisées. Alors que la plupart des études sur la perception multisensorielle des mouvements utilisent des stimuli discrets et spatialisés (Jack et Thurlow, 1973; Shams *et al.*, 2000; Kitagawa et Ichihara, 2002; Soto-Faraco *et al.*, 2003; Hidaka *et al.*, 2011), les sons qui ont été utilisés ici sont des sons continus et monophoniques, évoquant une dynamique à travers leurs variations de timbres.

Couplage visuo-moteur. L'expérience 9 a confirmé que les mouvements biologiques contraignaient effectivement la perception et la reproduction motrice d'un mouvement circulaire ou elliptique. Ce résultat confirme ceux de Viviani et ses collègues (Viviani et Stucchi, 1989; Viviani *et al.*, 1987; Viviani et Mounoud, 1990; Viviani *et al.*, 1997). Ces études ont suggéré que la perception des mouvements biologiques était traitée à un niveau amodal, c'est-à-dire à un niveau indépendant d'une modalité spécifique, en particulier pour la modalité visuelle et proprioceptive. Cependant, il est important de noter que les résultats mis en évidence par Viviani concernant le couplage visuo-moteur n'ont pas été réalisés en boucle ouverte visuelle (Viviani *et al.*, 1987; Viviani et Mounoud, 1990). Cela pouvant éventuellement laisser penser qu'un traitement spécifique entre la modalité visuelle et la production motrice existait. Ici, il a donc été mis en évidence qu'un cercle parcouru avec une vitesse non-biologique, i.e. accélérant dans les parties opposées, était perçu et reproduit plus aplati que la trajectoire réellement parcourue par le point lumineux. Inversement, une ellipse parcourue avec une vitesse non-biologique, i.e. accélérant dans les parties où sa courbure est maximale, est perçue et reproduite plus arrondie. Les résultats de notre expérience confirment donc que la relation sensori-motrice entre un mouvement visuel, biologique ou non, et sa reproduction en boucle ouverte semble bien être encodée sans traitement spécifique. Au-delà du couplage visuo-moteur, nous avons fait

l'hypothèse que la reproduction motrice devait traduire la perception visuelle du mouvement. Afin d'analyser les résultats dans une perspective plus perceptive, il est nécessaire de prendre en compte les contraintes biomécaniques et les synergies sous-jacentes à de tels mouvements.

La coordination motrice en jeu dans la production d'un mouvement graphique peut en effet être modélisée comme le couplage entre deux oscillateurs harmoniques amortis (cf. Chapitre 1). Le comportement physique d'un tel système donne lieu à l'existence d'attracteurs, états de coordination stables parmi l'ensemble des états possibles (Haken *et al.*, 1985; Zanone et Kelso, 1992; Kelso, 2012). De plus dans le cas spécifique de mouvements clos et répétitifs comme les mouvements mis en jeu dans nos expériences, des études ont montré que nous avons tendance à naturellement dessiner une ellipse plutôt qu'un cercle (Dounskaia *et al.*, 2000; Athenes *et al.*, 2004), et en particulier une ellipse d'excentricité .9. Sallagoity *et al.* (2004) ont en effet montré que si l'on demande à un sujet de dessiner un cercle le plus vite possible, il finissait par dessiner une ellipse d'excentricité .9. Par ailleurs, Danna *et al.* (2011) ont montré que cette excentricité était une des excentricités reproduite de façon préférentielle parmi l'ensemble des excentricités possibles. Au regard de ces considérations dynamiques, on peut ainsi estimer que la distorsion de la reproduction dans le cas des cercles parcourus avec une vitesse non biologique a été plus importante que dans le cas des ellipses car on a une plus grande facilité pour dessiner une ellipse qu'un cercle. Dans le cas des ellipses, bien que la distorsion géométrique soit moins grande que dans le cas des cercles, l'analyse de la distorsion cinématique montre que la cinématique des mouvements non-biologiques a été plus difficile à reproduire. Ce résultat est très intéressant car il met en évidence que l'on peut réussir à reproduire la géométrie d'un mouvement qui respecte une cinématique non naturelle. Par ailleurs, Wamain *et al.* (2011) ont récemment mis en évidence dans des expériences psychophysiques que nous discriminions plus aisément les ellipses statiques d'excentricités proches de celles que l'on dessine préférentiellement (i.e. avec une excentricité proche .9). On peut donc imaginer que ces ellipses sont également reproduites avec plus de facilité, même dans des situations cinématiques complexes.

Couplage audio-visuo-moteur. Dans la tâche de synchronisation audio-visuo-motrice, des sons de frottement continus ont été combinés avec les stimuli visuels de l'expérience 9 et il a été mis en évidence que ces stimuli sonores modifiaient substantiellement la reproduction du mouvement visuel. Comme dans le cas de l'expérience visuelle, les effets ont été plus importants pour les cercles que pour les ellipses. Les cercles parcourus à vitesse constante avec un son évoquant une ellipse ont été reproduits de façon significativement plus plate. Dans cette situation, les sons ont eu la même influence sur la reproduction que les cinématiques non biologiques dans le cas de l'expérience visuelle. Et de plus, dans les situations où la cinématique était déjà biaisée, i.e. non biologique, ces effets ont été encore plus forts, la corrélation entre les cinématiques non naturelles et les variations de timbre des sons évoquant des accélérations ont amplifié la perception visuelle des accélérations. Cet effet est à mettre en regard des connaissances sur l'importance de la corrélation temporelle entre les informations acoustiques et visuelles dans les processus d'intégration audiovisuelle. Parise *et al.* (2012) ont mis en évidence dans une tâche de localisation spatiale qu'un stimulus visuel était localisé de façon encore plus précise quand un son provenait du même endroit, ici, la corrélation entre les variations de la vitesse visuelle et celles évoquées par les variations de timbre des sons ont pu amplifier l'intégration de la vitesse visuelle qui a été traduite de façon encore plus saillante dans la reproduction du mouvement. Du point de vue temporel, la distorsion cinématique confirme que les sons incongruents avec la cinématique visuelle ont été plus difficilement reproduits. Et inversement, quand les sons étaient congruents avec la cinématique visuelle, biologique ou non, l'ajout de la cinématique sonore a amélioré la qualité de la reproduction. Les sons communiquent donc clai-

rement les variations de vitesse qui peuvent être ainsi plus facilement reproduites que quand elles sont transmises uniquement par l'entrée visuelle. A l'inverse des cercles, les ellipses ont été reproduites de façon significativement plus circulaires quand les sons évoquaient une vitesse constante. Cependant, les résultats ont montré que l'excentricité des ellipses semble avoir été moins affectée par l'ajout d'une cinématique sonore incongruente que l'excentricité des cercles. Cela peut être analysé de la même façon que pour de l'expérience purement visuelle. Les ellipses étant des formes dessinées de façon préférentielle (Athenes *et al.*, 2004) aux cercles, l'ajout d'une cinématique sonore incongruente a perturbé le mouvement mais pas autant que dans le cas des cercles qui sont moins fidèlement reproduits que des ellipses. De façon surprenante, la situation $e_{vk} = -.9$ et $e_{ak} = -.9$, c'est-à-dire la situation où le point parcourt un mouvement elliptique avec des accélérations sonores et visuelles cohérentes, n'a pas donné lieu à une reproduction motrice plus arrondie que dans le cas biologique ($e_{vk} = .9$, $e_{ak} = .9$), comme cela a été le cas dans la situation purement visuelle. Outre le fait que l'attracteur moteur semble avoir été suffisamment puissant pour contraindre la géométrie du mouvement bien que la cinématique visuelle était non naturelle, on peut faire l'hypothèse que ce stimulus audiovisuel a été intégré selon d'autres lois incarnées à haut niveau. Des études ont en effet mis en évidence que la loi en puissance $1/3$ n'était pas la seule loi qui contraignait la perception visuelle (Piaget et Matalon, 1958; Runeson, 1974) mais que l'on intégrait également des lois physiques comme les lois de Newton (La Scaleia *et al.*, 2014). L'incarnation de ces lois à travers le son a donc pu être traduite de façon différente dans les reproductions motrices, par exemple à travers les variations de la pression du stylet sur la tablette graphique, ces différences n'ayant pas eu nécessairement de conséquences sur l'excentricité de la forme et sur la distorsion cinématique.

L'intégration des mouvements biologiques à un niveau amodal a déjà été discutée dans le cas de l'expérience visuelle. L'expérience audiovisuelle montre donc que les mouvements biologiques sont effectivement intégrés à un niveau amodal et que la présentation d'informations incongruentes entre la modalité visuelle et la modalité auditive peut induire de nouvelles illusions. Du point de vue perceptif, on a vu dans le chapitre 2 que nous étions capables de reconnaître les mouvements biologiques du point de vue sonore. Dans l'expérience réalisée ici, on voit que des informations émanant de différentes modalités à propos des mouvements biologiques sont traitées comme un seul percept, c'est-à-dire que le système perceptif choisit l'interprétation la plus plausible des percepts incongrus pour les rendre cohérents du point de vue audiovisuel. Par exemple, les cercles parcourus à vitesse visuelle constante avec des sons évoquant des ellipses sont aplatis et la cinématique évoquée par le son joue le même rôle qu'une cinématique non-biologique du point de vue visuel.

Par dessous tout, cette expérience met en évidence une nouvelle situation où le son prend le dessus sur la modalité visuelle dans une tâche de reproduction d'un mouvement audiovisuel. Jusqu'à présent, la modalité auditive a souvent été considérée comme moins précise que la modalité visuelle dans la perception du mouvement. Cependant, la plupart des études ont placé la modalité auditive dans des situations où dès le départ, il était presque évident qu'elle ne pourrait pas influencer la modalité visuelle. En effet, la plupart de ces études ont étudié la perception audiovisuelle du mouvement avec des stimuli sonores discrets et spatialisés *imitant* ainsi les stimuli visuels. Dans ces situations, l'audition est nécessairement moins précise que la vision car on localise naturellement moins bien un stimulus sonore qu'un flash lumineux. Seules quelques études semblent avoir vraiment mis en évidence que l'audition pouvait prendre le dessus sur la vision. Shams *et al.* (2000) ont montré que la perception du nombre de flashes lumineux générés de façon très rapprochée était modifiée si on ajoutait des *bips* sonores en nombre différents. Concernant le mouvement, Sekuler et Lau (1997) ont montré que la perception visuelle du mouvement de deux billes qui se croisent était mo-

difiée si on ajoutait un *bip* au moment où les deux billes se rencontrent : sans le *bip* on a l'impression que les billes se croisent, avec le *bip* on a l'impression qu'elles rebondissent l'une sur l'autre.

Nos résultats révèlent le rôle des interactions perceptivo-motrices dans l'intégration multisensorielle du mouvement. Ce paradigme ouvre de nombreuses perspectives qui seront discutées plus en détail dans le chapitre 6, on peut néanmoins mentionner qu'il ouvre des perspectives pour la compréhension des processus cognitifs impliqués dans la perception audiovisuelle des mouvements biologiques (Grossman *et al.*, 2000; Grezes *et al.*, 2001; Giese et Poggio, 2003; Dayan *et al.*, 2007) et comment nous faisons émerger un monde cohérent à partir de nos contraintes sensori-motrices (O'Regan et Noë, 2001; Casile et Giese, 2006; Thornton et Knoblich, 2006). De plus, ce paradigme offre un nouveau contexte d'études pour les recherches concernant les problèmes neurologiques comme l'autisme et la schizophrénie, impliquant des dysfonctionnements perceptifs au niveau visuel (Freitag *et al.*, 2008; Klin *et al.*, 2009) ou au niveau auditif (Micoulaud-Franchi *et al.*, 2011, 2012).

Chapitre 5

Vers la création de métaphores sonores et la sonification du geste d'écriture


Sommaire

5.1	Création de métaphores sonores évoquant un geste humain	126
5.2	Vers l'apprentissage d'un geste expert avec le son : une nouvelle écriture	128
5.3	Diagnostic et réhabilitation de la dysgraphie	136
5.4	Conclusion	137

La première partie de cette thèse a mis en évidence des aspects fondamentaux sur les relations entre les variations de timbre d'un son de frottement, la façon dont on le perçoit, les caractéristiques du geste qu'il évoque, et enfin son rôle dans des tâches de synchronisation visuo-motrice et audio-visuo-motrice. En particulier, il a été mis en évidence que les variations de timbre des sons de frottement évoquaient la cinématique du mouvement et que cette cinématique évoquée pouvait être associée à des formes (cf. Chapitre 2). On a également vu que la relation entre la forme visuelle évoquée et la forme transmise par les sons n'était pas bijective et qu'une distorsion existait entre la *forme visuelle* et la *forme sonore* (cf. chapitre 3). Cette distorsion a révélé que la forme imaginée à partir du son était plus aplatie que la forme originale. Enfin, dans une tâche de synchronisation entre un mouvement et des stimuli audiovisuels, on a vu que la modalité sonore transmettait la dynamique d'un mouvement de façon beaucoup plus précise que la modalité visuelle. Tous ces résultats ouvrent de nombreuses perspectives d'application, d'une part la création de sons inouïs, appelés métaphores sonores, évoquant des événements sonores qui n'existent pas dans la réalité mais qui évoquent malgré tout une action sur l'objet. Par ailleurs, on peut également envisager la création d'outils d'apprentissage ou de réhabilitation motrice basés sur un retour sonore. Ces perspectives ont été développées dans le cadre du projet projet MetaSon qui s'articule autour de deux problématiques. La première concerne la sonification des voitures électriques. Ces voitures sont effectivement très silencieuses et n'émettent aucun bruit de moteur permettant d'informer les piétons de leur présence et la nécessité de se déplacer pour éviter un accident. Par ailleurs, le fait qu'elles n'émettent aucun son est très perturbant pour les conducteurs. En effet, sans que nous en soyons conscient, les voitures thermiques classiques émettent des sons qui nous informent très précisément sur notre conduite. Le premier objectif du projet est donc de sonifier le mouvement de ces voitures, d'une part à l'extérieur pour informer les piétons de la présence d'un véhicule, et d'autre part à l'intérieur, pour informer les conducteurs de la dynamique du véhicule et sur leur mode de conduite. Cette tâche nécessite donc d'une part de

déterminer les variables à sonifier (vitesse ? accélération ? ...) et d'autre part de sélectionner les sons qui informent le mieux les piétons et les conducteurs sur l'état du véhicule. La seconde application du projet qui nous intéresse plus particulièrement, concerne la sonification de l'écriture pour la détection et la réhabilitation d'une pathologie graphomotrice, la dysgraphie, cette problématique nécessite donc de mettre en place une stratégie de sonification des paramètres dynamiques du geste pour informer une personne en temps-réel sur la qualité de son geste d'écriture.

5.1 Création de métaphores sonores évoquant un geste humain

 Les études présentées dans les chapitres précédents nous ont permis d'identifier les invariants perceptifs relatifs à l'évocation d'un geste humain qui frotte sur une surface. On sait désormais créer des sons qui contiennent une information acoustique relative à la reconnaissance auditive d'un geste. Par ailleurs, le paradigme action → objet tel que nous l'avons présenté dans le chapitre 1 dans le contexte du contrôle intuitif de la synthèse permet d'envisager la création de sons inouïs que l'on pourrait appeler métaphores sonores en associant des actions et des objets physiquement incongrus. Cette possibilité est envisageable dans la mesure où l'implémentation source-filtre permet de croiser les informations relatives à l'action, contenues essentiellement dans la source, et les informations relatives à l'objet, contenues essentiellement dans le filtre. Peut-on par exemple créer le son d'une personne en train de frotter sur du vent ? En utilisant les connaissances sur le geste, on peut même envisager de frotter le vent de différentes façons, par exemple le frotter de façon fluide ou bien saccadé.

En linguistique, une métaphore correspond à une transformation du signifié, i.e. le sens d'un mot, par un usage du signifiant, i.e. le signal acoustique de parole correspondant au mot, dans un contexte différent de celui dans lequel il est utilisé habituellement¹. Ainsi, une métaphore sonore pourrait être définie comme la création d'un son évoquant une action particulière, comme frotter de façon fluide, sur un objet que l'on ne peut *a priori* pas frotter, par exemple le vent ou l'eau (voir figure 5.1).

Pour la génération de l'action, qui est essentiellement caractérisée par le signal source, on peut s'appuyer sur les résultats du chapitre 2 qui montrent qu'en appliquant un filtrage passe-bas contrôlé par des profils de vitesse particuliers, on peut évoquer différents types de gestes, et en particulier des gestes fluides. En modifiant la valeur de l'exposant de la loi en puissance reliant la vitesse à la courbure de la forme, on peut de plus envisager de changer le type de geste et le rendre par exemple plus saccadé.

Pour la contribution de l'objet, différentes façons peuvent être envisagées pour évoquer des objets liquides ou aériens. Pour les liquides par exemple, il est possible de remplacer le résonateur solide par un résonateur correspondant à une goutte d'eau que l'on peut synthétiser très simplement (van den Doel, 2005) et ensuite générer le son évoquant une personne qui frotte de l'eau en appliquant le même processus de synthèse que pour le cas des solides. La contribution de l'objet se trouve également dans le bruit utilisé en entrée du modèle de synthèse, en effet, du point de vue solide, ce bruit correspond à la surface de l'objet vue par l'excitateur. On peut donc également envisager de remplacer le bruit, correspondant à la surface vue par l'objet frottant dans le cas solide par

1. Plusieurs définitions et interprétations de la notion de signifié et signifiant ont été proposées en linguistique par Ferdinand De Saussure (1916) puis en psychanalyse par Jacques Lacan (1966), ici nous référons à ces notions dans l'unique but de pointer la différence entre le sens d'un mot ou d'un son, et le signal acoustique en lui-même. Bien qu'il pourraient être intéressant de les envisager, les enjeux et perspectives linguistiques (et éventuellement psychanalytiques) dépassent largement l'ambition de cette thèse.

une texture stationnaire d'eau qui coule et ensuite lui appliquer le processus de filtrage passe-bas ou passe-bande reliée à la vitesse du geste pour évoquer le fait de frotter sur une surface liquide. Pour les sons aérodynamiques, (Farnell, 2010; Verron *et al.*, 2010) ont proposé de modéliser la dynamique interne du vent par un banc de filtres passe-bande évolutifs permettant d'évoquer soit un souffle constant soit des rafales. Ici on peut ainsi envisager de contrôler ce banc de filtres résonnants par une dynamique gestuelle afin d'évoquer le son d'une personne qui frotte le vent.

La création de ces métaphores permet ainsi d'évaluer la robustesse des invariants perceptifs et plus particulièrement si l'on est capable de reconnaître un geste à travers ces sons inouïs. Une expérience formelle serait nécessaire pour valider cette affirmation. Ces métaphores sonores ont par ailleurs des applications évidentes dans le design sonore et la musique électroacoustique. En outre, elles pourraient être utilisées comme stratégie de sonification des gestes du conducteur pour les informer des mouvements qu'ils réalisent. En effet, le contrôle de la génération de ces sons inouïs directement à partir de descripteurs sémantiques simples permet d'envisager une utilisation par des utilisateurs naïfs, ne connaissant a priori rien au processus de synthèse sous-jacent.

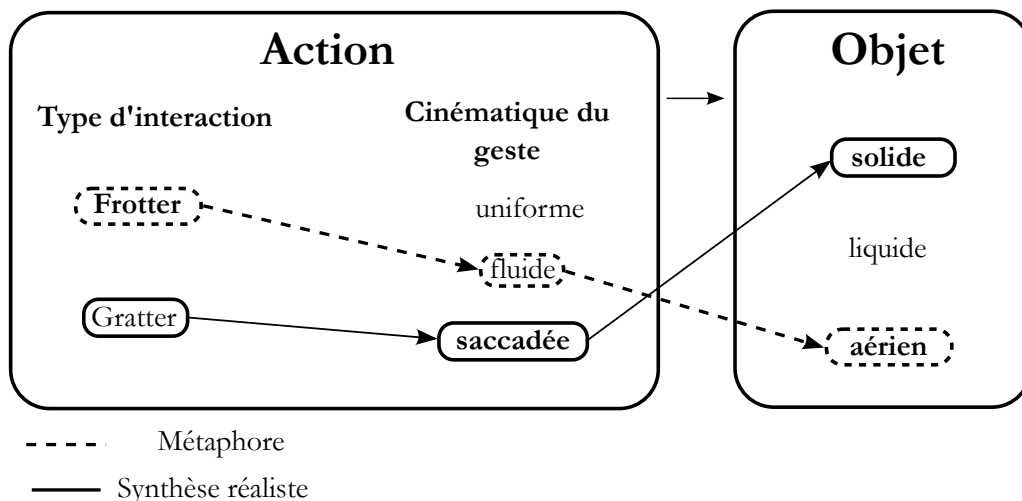


FIGURE 5.1 – Processus de création de métaphores sonores basé sur le croisement entre des invariants reliés au type d'interaction (frotter – gratter), à la cinématique du geste (uniforme – fluide – saccadé) et ceux liés à l'objet (solide – liquide – aérien). On voit que ce concept permet à la fois de créer des situations réalistes (flèches pleines), *Gratter de façon saccadée sur un résonateur solide*, mais également des métaphores sonores (flèches en pointillés), *Frotter de façon fluide sur un résonateur aérien, i.e. frotter le vent*.

Dans la lignée de cette idée de contrôler la synthèse du son par des moyen intuitifs, on peut envisager de contrôler la génération du mouvement par des dessins. Etant donné que le profil de vitesse d'un geste peut être généré à partir d'une forme géométrique, et plus particulièrement de la courbure locale en appliquant la loi en puissance $1/3$, il est possible de recréer virtuellement l'étape de production du dessin et ainsi de générer le son qui serait produit par le scripteur en train de dessiner cette forme. Ainsi en utilisant des textures abstraites comme celles utilisées par Adrien Merer (2011) pour créer des synthétiseurs de sons évoquant des mouvements avec un contrôle intuitif, il semble possible de proposer des contrôle intuitifs basés sur une description graphique du son qui va être généré. Dans les études d'Adrien Merer, le mouvement était considéré de façon beaucoup plus générale et ne concernait pas spécifiquement les mouvements humains, nos résultats permettraient de raffiner les contrôles déjà proposés par rapport à cette catégorie de mouvements particuliers.

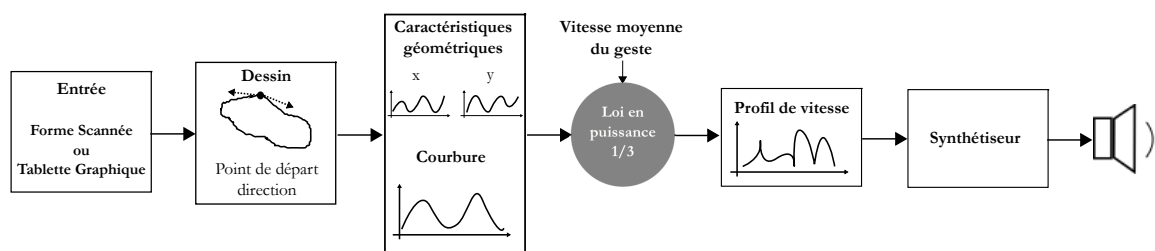


FIGURE 5.2 – Sonification d’une forme géométrique – figure issue de Thoret *et al.* (2012a)

5.2 Vers l’apprentissage d’un geste expert avec le son : une nouvelle écriture


La section précédente a mis en évidence une première application des travaux présentés dans cette thèse pour le design et le contrôle des synthétiseurs. Dans cette section, nous allons présenter une seconde application qui concerne l’apprentissage d’un geste avec pour support la modalité auditive. Plus particulièrement, on souhaite évaluer si un retour sonore permettrait d’apprendre plus facilement un nouveau geste d’écriture. Pour sonifier un geste et plus particulièrement un geste d’écriture, il est nécessaire de définir un geste de référence vers lequel un geste doit tendre pour atteindre une bonne qualité. Cette référence est définie par rapport à des variables caractéristiques du geste d’écriture qui vont être présentées.

5.2.1 Variables caractéristiques du geste d’écriture

Une étape nécessaire pour mettre en place une stratégie de sonification et d’apprentissage d’un geste par le biais du son est de définir les variables les plus caractéristiques de ce type de geste. Dans les chapitres 2, 3, et 4, on a mis en évidence que le profil de vitesse d’un geste était caractéristique du geste du point de vue sonore, c’est donc une variable importante à considérer pour la sonification d’un geste d’écriture.

Par ailleurs, une étude menée par Danna *et al.* (2013a) a permis d’établir un descripteur caractérisant la qualité de l’écriture chez les enfants dysgraphiques (nous nous intéresserons plus tard à cette pathologie). Ce descripteur caractérise la *disfluente* (*disfluency*) du mouvement d’écriture. Plus particulièrement, un *mauvais* geste d’écriture peut être caractérisé par un mouvement très peu fluide qui présente des pics de vitesse en quantité très supérieure à celui d’un *bon* geste d’écriture (voir figure 5.3). En pratique, le nombre de pics surnuméraires correspond à la différence de pics de vitesse entre le profil de vitesse filtré passe-bas² à 10 Hz et celui filtré à 5 Hz (voir figure 5.3). Danna *et al.* (2013a) ont alors pu observer que les profils de vitesse des enfants qui présentent des troubles de l’écriture présentaient un nombre de pics surnuméraires significativement supérieur à ceux des enfants normo-scripteurs. Cette variable caractérise donc bien la qualité d’un geste d’écriture et peut donc être intégrée dans un processus de sonification d’un geste d’écriture.

5.2.2 Stratégie de sonification : la métaphore du couinement

 La stratégie de sonification vise donc à informer le scripteur de la qualité de son geste, en particulier si une vitesse minimale d’exécution est atteinte ou pas (ayant pour conséquence une plus grande fluidité du mouvement). Nous avons ainsi choisi d’illustrer cette distinction par des retours sonores différents : lorsque le scripteur produit un bon geste d’écriture (avec une vitesse correcte), un son de frottement (comme ceux utilisés dans les chapitres précédents) est généré. En revanche,

2. Le filtrage passe-bas a l’effet d’un *lissage*.

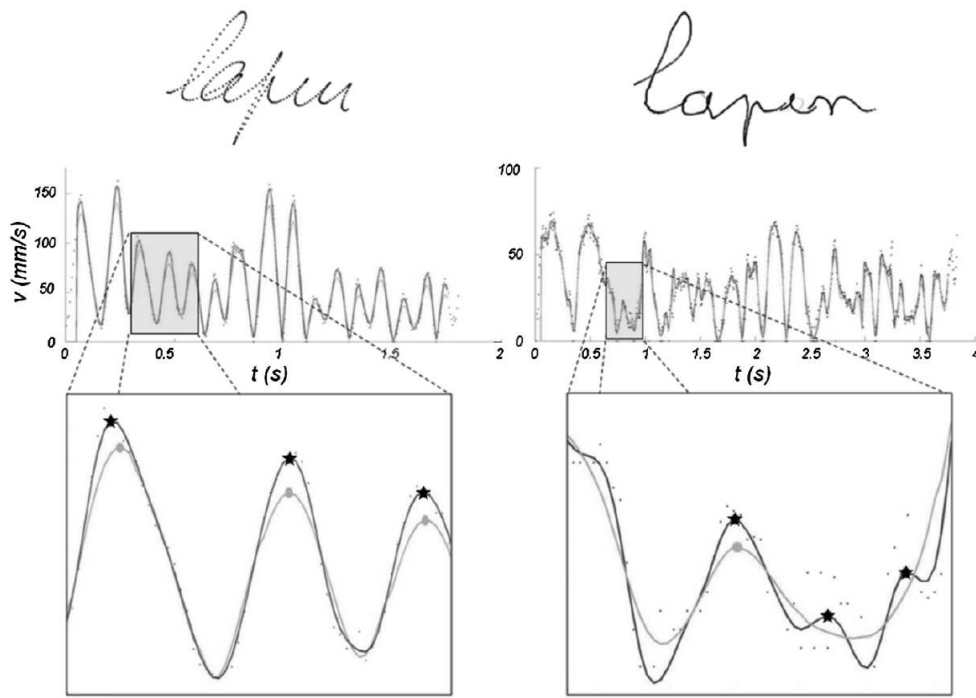


FIGURE 5.3 – Exemples de productions graphiques et de profils de vitesse associés à un bon geste d'écriture (gauche) et à un mauvais geste d'écriture (droite) (d'après Danna *et al.* (2013a)) chez les enfants normo-scripteurs et dysgraphiques. On voit clairement que le profil de vitesse à un mauvais geste a plus de pics de vitesse surnuméraires que celui associé à un mauvais geste. Les points correspondent aux pics de vitesse des profils.

lorsqu'il est trop lent, un son de couinement, souvent ressenti comme un son plutôt désagréable, et évoquant par exemple au son d'une mine qui crisse, est généré. Nous avons proposé une méthode pour synthétiser ces sons et les contrôler intuitivement en accord avec le paradigme action → objet (Thoret *et al.*, b), cette méthode va être rappelée dans la section suivante.

Considérations physiques : modèle de friction de Coulomb

Les sons de couinements sont générés par des phénomènes physiques différents de ceux impliqués dans la génération des sons de frottement bruités utilisés dans les chapitres précédents. En particulier, ils mettent en jeu des phénomènes non linéaires.

La friction non linéaire est un sujet de recherche historique étudié dès le 18ème siècle par Coulomb qui a donné son nom à un cas particulier de friction. En première approximation, la friction de Coulomb peut être décrite très simplement par un modèle phénoménologique aussi appelé *modèle du tapis roulant* qui est schématisé en figure 5.4. Quand deux objets sont en contact frottant et quand

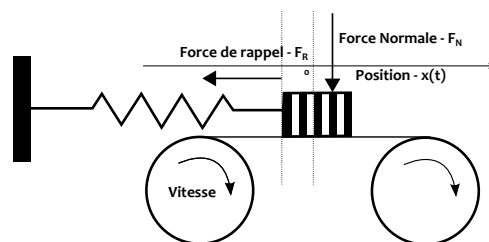


FIGURE 5.4 – Modèle du tapis roulant

la force d'interaction entre les deux objets devient suffisamment grande, le phénomène physique peut être décrit comme le déplacement d'une masse retenue par un ressort sur un tapis roulant qui

la déplace continûment de sa position d'équilibre à une vitesse v . Quand la force de friction F_{fr} (i.e. la pression entre les deux corps) devient plus faible que la force la force résultante de la force normale F_N et de la force de rappel F_r du ressort, la masse dérape dans le sens opposé jusqu'à ce que le rapport des forces se ré-inverse. Cette simplification peut être décrite très simplement : la masse est déplacée linéairement de sa position d'équilibre et y revient instantanément dès qu'elle atteint la phase de glissement. Cette modélisation décrit simplement les phénomènes de friction non linéaire, elle est appelée fréquemment mouvement *collé-glissé* (stick-slip en anglais) ou encore mouvement de Helmholtz. Un exemple de ce type de mouvement est présenté en figure 5.5. Si on néglige la phase de

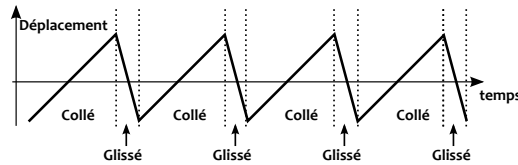


FIGURE 5.5 – Un cas particulier de mouvement de Helmholtz (collé-glissé)

glissement, le mouvement résultant $x(t)$ de la masse sur le tapis roulant correspond à un signal en dents de scie dont la décomposition en série de Fourier s'écrit :

$$x(t) = \sum_{k=1} (-1)^k \frac{\sin(\omega_0 k t)}{k} \quad (5.1)$$

où ω_0 correspond à la fréquence des phases où la masse est collée au tapis roulant et k correspond à l'ordre des différentes harmoniques. La dent de scie est la description la plus simple des modèles de friction non linéaire.

Ce modèle nous donne un premier *a priori* sur le comportement de la source (i.e. de l'excitation) dans l'implémentation source-filtre souhaitée : la source doit avoir un spectre harmonique. Cette modélisation nous permet donc d'avoir une première idée de l'invariant transformationnel associé au phénomène de friction non linéaire bien que nous soyons conscients que ce type de modélisation ne soit pas généralisable à tous les phénomènes de friction non linéaire.

Considérations acoustiques : observations sur des sons enregistrés

Des observations empiriques de sons de couinement produits par un doigt humide sur un verre, et de sons harmoniques produits par un doigt faisant *chanter* un verre vont être présentées. Tous les sons ont été enregistrés en chambre anéchoïque avec un microphone Neumann-KM84i cardioïde positionné à 30 cm au dessus du verre. Une analyse de l'harmonicité du spectre sera réalisée pour chaque enregistrement. L'analyse de l'harmonicité du spectre a été réalisée en utilisant une toolbox développée par Dan Ellis qui permet la détection temporelle des pics d'un spectre d'un extrait sonore (Ellis, 2003). Une régression linéaire sur la valeur de l'inharmonicité des spectres a été réalisée, $\gamma_n = \frac{f_n}{nf_0} - 1$, où f_n est la fréquence du n -ième pic détecté et où f_0 est la fréquence fondamentale du spectre. Si l'ensemble des partiels suit bien une loi harmonique, γ_n doit être nul. Nous présentons ici nos observations sur l'exemple du doigt humide sur un verre qui couine et chante car il illustre bien les différents comportements non linéaires qui nous intéressent. D'autres exemples sont présentés en annexe 7.5 (porte qui grince, assiette qui couine).

La figure 5.6 présente la transformée de Gabor de l'enregistrement d'une personne qui fait couiner et chanter un verre humide. Quand le bord d'un verre est frotté, le son produit présente différents comportements. Le premier d'entre eux correspond à une variation de la fréquence fondamentale f_0 de façon chaotique autour d'une valeur moyenne. Le second type de comportement, quand le verre

chante, se produit quand la valeur de f_0 se bloque un mode du verre. Les transitions entre ces deux états semblent difficilement prévisibles.

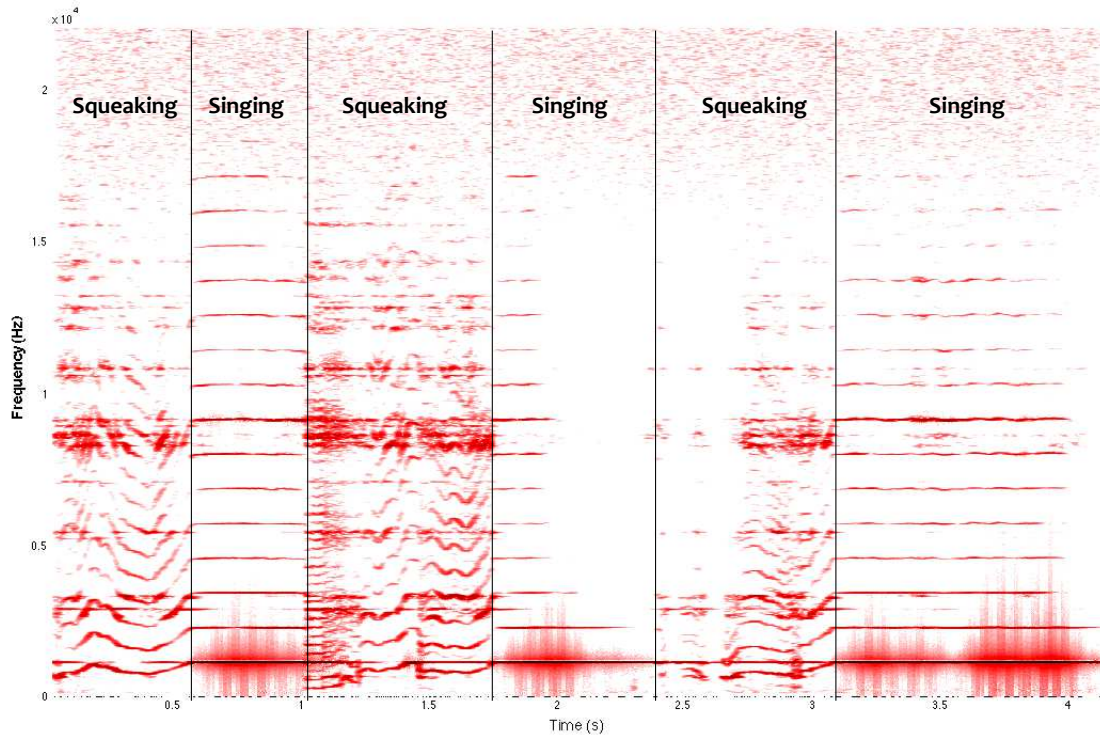


FIGURE 5.6 – Transformée de Gabor du son produit par un doigt qui fait couiner et chanter un verre – Figure issue de Thoret *et al.* (b)

A partir de ces observations empiriques, on peut faire l'hypothèse que la morphologie acoustique qui caractérise tous ces comportements est l'évolution de la fréquence fondamentale f_0 du spectre harmonique au cours du temps. Ces différents comportements peuvent donc être modélisés par un spectre harmonique dont la fréquence fondamentale varie au cours du temps. Cela est bien connu pour les instruments de musique dont la modélisation physique est également un couplage non linéaire entre un exciteur et un résonateur.

Modèle de synthèse de friction non linéaire

Pour créer notre modèle de synthèse de sons de friction non linéaire, on souhaite conserver le paradigme action → objet basé sur les invariants perceptifs. Cependant, nous venons de voir que le phénomène de friction non linéaire induit un couplage entre l'exciteur et le résonateur, ce qui est incompatible dans la perspective d'une implémentation source-filtre. Cependant, notre but ici n'étant pas de reproduire le comportement physique de ce type de situation mais de générer des sons évoquant ces actions du point de vue perceptif. Ainsi, tout en conservant le paradigme action → objet, nous proposons de générer un signal d'excitation *équivalent* permettant une synthèse correcte des morphologies sonores associées. Ce problème est schématisé par la figure 5.7.

A partir des considérations physiques et spectrales présentées dans les sections précédentes, le modèle de synthèse général suivant a été proposé pour modéliser l'excitation $e(t)$ du système source résonance :

$$e(t) = \sum_{k=1} AM_k(t) \sin(\Omega_k(t)) \quad (5.2)$$

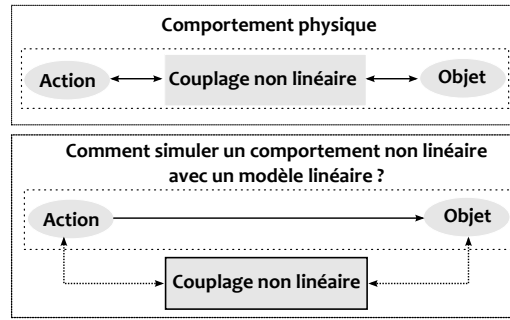


FIGURE 5.7 – Comment séparer le couplage non-linéaire afin d’implémenter des comportements non linéaires avec un filtrage linéaire ?

où la fréquence instantanée est définie par :

$$\frac{d\Omega_k}{dt}(t) = 2\pi k f_0(t) \quad (5.3)$$

et AM_k est la modulation d’amplitude appliquée à l’harmonique d’ordre k .

- **Evolution de la fréquence instantanée.** Dans le monde qui nous entoure, on peut faire l’hypothèse qu’un son de friction non linéaire est principalement contrôlé par la pression $p(t)$ de l’excitateur et sa vitesse relative $v(t)$ par rapport à l’objet frotté. Du point de vue du signal acoustique, les morphologies sonores caractérisant les grincements, les couinements, ou encore les auto-oscillations (les moments où le verre *chante*) sont différenciées uniquement par l’évolution temporelle $f_0(t)$ de la fréquence fondamentale. Ainsi, la formule suivante a été proposée pour définir l’ensemble des comportement possibles de f_0 :

$$f_0(t) = (1 + \epsilon(t)\gamma)\Phi(v(t), p(t)) \quad (5.4)$$

où Φ est une fonction qui relie les paramètres dynamiques de pression $p(t)$ et de vitesse $v(t)$ à la fréquence fondamentale du spectre. Dans le cas des auto-oscillations, la fréquence fondamentale est bloquée sur un mode de l’objet en vibration. ϵ correspond à un bruit filtré passe-bas avec une fréquence de coupure inférieure à 20 Hz. Cette variable stochastique permet de modéliser les variations chaotiques observées dans le cas des sons de couinement et de grincement. Dans le cas des auto-oscillations, ϵ est nul. γ peut être défini comme un paramètre définissant l’état du système et vaut 1 pour les grincement ou les couinements et 0 dans le cas des auto-oscillations. La formulation présentée ici considère que Φ est reliée au couple pression/vitesse. Cependant, dans une perspective d’application à la sonification, il est important de noter que Φ peut être défini par rapport à n’importe quel descripteur dynamique en fonction de l’application considérée comme cela a été mentionné plus haut dans ce chapitre.

- **Modulation d’amplitude.** Dans le cas particulier du verre qui chante, des modulations d’amplitudes de périodes très lentes ont été observées dans les enregistrements. Rossing (1994) a mis en évidence dans une étude portant spécifiquement sur les verres que cette modulation d’amplitude était directement reliée à la vitesse $v(t)$ du doigt humide en train de frotter le bord du verre et au diamètre D du verre. Cette modulation s’exprime sous la forme suivante :

$$AM_k(t) = \frac{1}{k} \sin \left(2\pi \int_0^t \frac{v(t)}{\pi D} dt \right) \quad (5.5)$$

Dit autrement, plus le diamètre du verre est petit et plus la vitesse du doigt est grande, plus la modulation d'amplitude sera rapide, et inversement. A part dans ce cas particulier du verre, aucune autre loi de modulation d'amplitude n'a été trouvée dans la littérature. Pour les autres comportements, en première approximation et en se basant sur les observation spectrales, l'amplitude de chaque partiel suit une décroissance en $1/k$: $AM_k = \frac{1}{k}$.

Finalement, cette section a présenté un modèle général permettant de simuler les différentes morphologies sonores reliées aux phénomènes de friction non linéaire que l'on souhaite modéliser. La table 5.1 résume les différents paramètres du modèle de synthèse en fonction des différents comportements.

TABLE 5.1 – Implémentation des différents comportements non linéaires

	$f_0(t) =$	$AM_k(t) =$
Porte qui grince Vaisselle qui couine Verre qui couine	$(1 + \epsilon(t))\Phi(\Gamma(t))$	$\frac{1}{k}$
Verre qui chante	$\tilde{f}_{(n,0)}$	$\frac{1}{k}\sin\left(2\pi\int_0^t\frac{v(t)}{\pi D}dt\right)$
Corde frottée	$\tilde{f}_{(n,0)}$	$\frac{1}{k}$

$\Gamma(t)$ représente le mapping entre les descripteurs, vitesse et pression dans une situation physique, et la fréquence fondamentale – $\tilde{f}_{(n,0)}$ est le n-ème mode de l'objet vibrant.

Stratégie de contrôle

Le modèle de synthèse proposé peut être contrôlé par des paramètres dynamiques comme la vitesse et la pression. Ce contrôle est proche de celui des instruments de musique de type auto-entretenus comme le violon où la clarinette. Dans le cas de notre étude, la pression sera supposée constante et seule la vitesse sera reliée au type de son généré. Des contrôles plus raffinés permettant par exemple ce modèle en fonction de la jouabilité d'un instrument à cordes frottées peuvent être envisagés, ces contrôles seront détaillés dans les perspectives de cette thèse en section 6.2.1.

5.2.3 Une preuve de concept : apprentissage d'un geste par la main non-dominante

Afin de valider cette stratégie de sonification pour l'apprentissage d'un geste d'écriture, nous allons présenter une expérience princeps lors de laquelle des sujets devaient apprendre des nouveaux caractères graphiques avec leur main non dominante. Les sujets apprenaient certains caractères avec la stratégie de sonification précédente et d'autres sans afin d'évaluer l'intérêt de la sonification pour l'apprentissage d'un geste d'écriture. Cette expérience est extraite de l'article suivant :

Danna, J., Fontaine, M., Paz-Villagrán, V., Gondre, Thoret, E., Aramaki, M., Kronland-Martinet, R., Ystad, S. Velay, J.-L. (In press) The Effect of Real-Time Auditory Feedback on Learning New Characters. *Human Movement Science*

Methodes. 32 sujets droitiers âgés d'en moyenne 25.2 ans ($SD = 3.5$) ont été répartis en deux groupes. Tous avaient une vision normale ou corrigée et aucun d'entre eux n'avait de problème d'audition. La tâche consistait à écrire sur une feuille de papier posée sur une tablette graphique les 4 caractères (voir figure 5.8) présentés au-dessus de la feuille. Le protocole expérimental était séparé en quatre parties, une courte phase de familiarisation, un pré-test, et deux post-tests (voir figure 5.9). Le premier test (pré-test) était réalisé juste après la tâche de familiarisation, et le second post-test le lendemain. Lors la phase de familiarisation, tous les participants ont pu écrire une fois chaque caractère

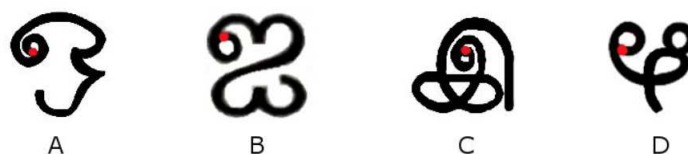


FIGURE 5.8 – Caractères que les sujets devaient reproduire – le point rouge indiquait la position où ils devaient commencer le mouvement (d'après Danna *et al.* (ress)).

sans retour sonore. Durant la phase d'entraînement, les sujets ont écrit les 4 caractères 16 fois chacun. Ils avaient un retour sonore seulement pour deux des quatre caractères. La différence entre les deux groupes concernait l'ordre dans lequel ils apprenaient à écrire les lettres avec le retour sonore lors de la phase d'entraînement. Plus précisément, le groupe 1 a commencé par apprendre les deux premiers caractères sans le retour sonore puis les deux seconds avec le retour sonore. Le groupe 2 a appris les deux premiers caractères tout d'abord avec le retour sonore puis les deux autres sans. L'ordre dans lequel les sujets ont appris les lettres a été pseudo-contrebalancé entre tous les participants de chaque groupe. La stratégie de sonification utilisée pour le retour sonore correspond à celle présentée dans

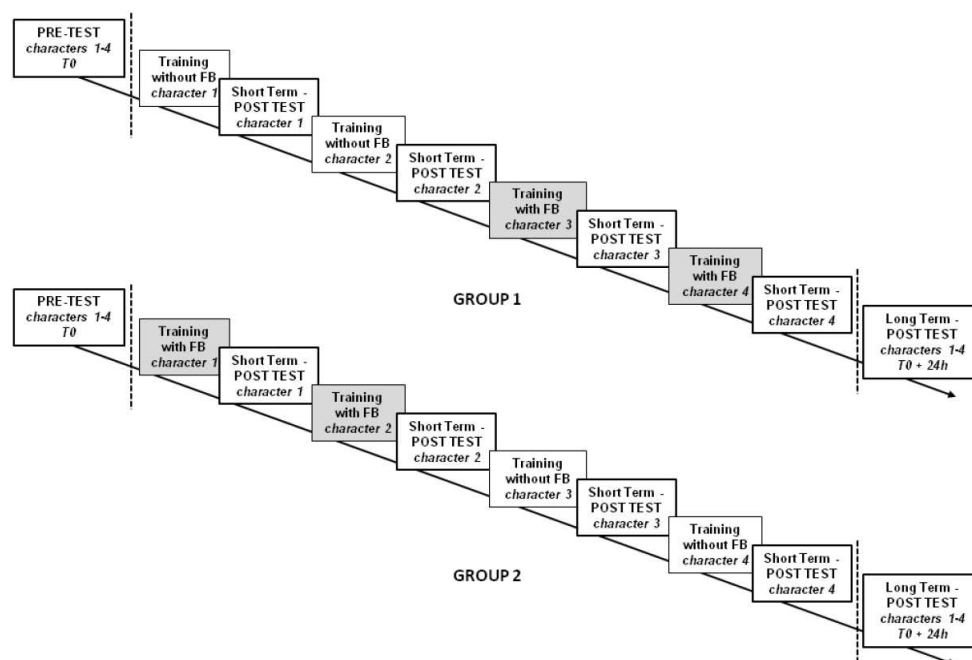



FIGURE 5.9 – Protocole expérimental pour les deux groupes de sujets (d'après Danna *et al.* (ress)).

la section précédente. Les pics surnuméraires caractérisant la disfluece ont également été sonifiés par des sons d'impacts. La sonification de la disfluece par des sons d'impacts n'a pas pour objectif de guider le geste d'écriture en temps réel en raison de son aspect discret, l'objectif de sonifier ces pics est de donner un retour global sur la qualité de l'écriture à l'enfant au fur et à mesure des essais :

plus son écriture est bonne, moins il doit y avoir de pics.

Analyses des données. Deux descripteurs cinématiques et deux descripteurs spatiaux ont été analysés. Les descripteurs cinématiques correspondent à la vitesse moyenne du mouvement réalisé pour produire un caractère et au nombre de pics surnuméraires (descripteur présenté précédemment). Les descripteurs spatiaux sont la taille de la forme (sa *longueur cumulée*) et une distance, pouvant être assimilée à la distance DTW (Dynamic Time Warping) entre la forme dessinée et la forme de référence. Dit autrement, la distance calculée permet de déterminer la distorsion entre la géométrie de la forme dessinée par les sujets et celle de référence. Il est important de noter qu'elle ne correspond pas exactement au Dynamic Time Warping défini mathématiquement par Keogh et Ratanamahatana (2005).

Analyse statistique. Chaque descripteur a été moyenné séparément pour les deux caractères réalisés dans la phase d'entraînement avec le retour sonore et ceux réalisés sans le retour sonore. Une ANOVA à 3 facteurs : 3 Sessions d'apprentissage (pré-test, post-test à court terme, post-test à long terme) x 2 Sonifications (avec ou sans sonification) x Ordre de présentation du retour sonore (Retour sonore en premier ou en second).

Résultats. Les résultats ont clairement montré un intérêt du son dans l'apprentissage de nouveaux caractères avec la main non-dominante. En effet, on a pu observer que quand les participants ont commencé par apprendre les caractères sans retour sonore, les caractères appris avec retour sonore étaient reproduits avec une vitesse moyenne significativement plus rapide et fluide que ceux appris sans retour sonore dans le post-test à court terme. Et de plus, pour le groupe qui a commencé à apprendre avec le retour sonore, les caractères appris sans le retour sonore ont été écrits avec la même vitesse moyenne et la même fluidité que ceux appris avec le retour sonore, montrant ainsi un *transfert de compétence* par le biais de l'apprentissage avec le son (voir 5.10(a)). Cela montre donc le même transfert de compétence que pour la vitesse moyenne (voir 5.10(b)). On doit pondérer ces deux premiers résultats par le fait que la précision spatiale, caractérisée par la distance DTW, est meilleure pour les caractères ayant été appris sans retour sonore. Enfin, la longueur des caractères n'a pas été affectée par la présence ou non d'un retour sonore. Une vidéo présentant un exemple de réalisation de la tâche est disponible sur la page web associée à la thèse. 

Discussion

Cette expérience met donc clairement en évidence l'intérêt d'un retour sonore dans une tâche d'apprentissage du geste d'écriture basé sur un critère de fluidité du mouvement. On voit que les caractères appris avec le retour sonore en premier sont mieux reproduits dans le post test à court terme, et que les caractères appris sans le retour-sonore pour le groupe de sujets ayant commencé par apprendre des caractères avec le retour sonore bénéficient de l'apprentissage avec le son pour les deux premiers. Ce transfert du retour sonore aux caractères appris sans l'utilisation du son est particulièrement intéressant et peut être discuté dans le contexte idéomoteur, aussi appelé théorie de l'encodage des événements (TEC : Theory of Event Coding) déjà évoquée dans le chapitre 2 (Hommel *et al.*, 2001; Prinz, 1997). Comme pour la loi en puissance 1/3, on peut faire l'hypothèse de l'existence d'un percept unifié pour le geste d'écriture qui, une fois appris, peut être réactivé par différentes modalités. L'apprentissage de certains caractères avec le son a donc pu bénéficier aux caractères appris sans le son, et en particulier au niveau des aspects dynamiques des mouvements. Etant donné

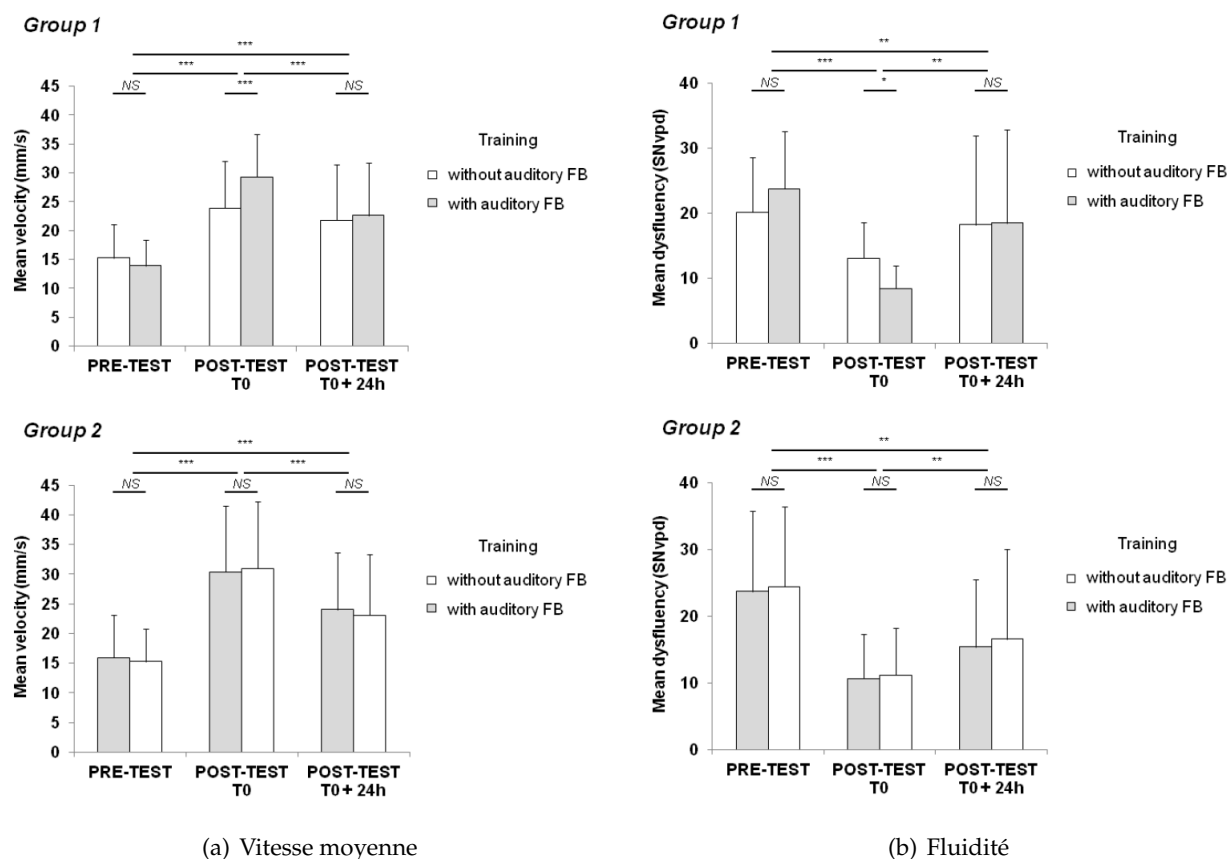


FIGURE 5.10 – Descripteurs des performances motrices pour les différents tests du plan d'expérience (d'après Danna *et al.* (ress))

que la précision spatiale était meilleure pour les formes apprises sans retour sonore, on peut faire l'hypothèse que la charge cognitive induite par la modalité auditive pour reproduire la dynamique adéquate a entraîné une diminution de l'attention portée par les sujets à la forme en elle même.

5.3 Diagnostique et réhabilitation de la dysgraphie

Les sections précédentes ont présenté différentes expériences et stratégies de synthèse permettant d'envisager d'utiliser le son comme outil de guidage et d'apprentissage des caractéristiques dynamiques d'un geste d'écriture. Ces différentes stratégies sont actuellement mises en application dans le cadre d'une collaboration avec des membres de l'équipe *Dynamique des apprentissages auditifs et moteurs* du Laboratoire de Neurosciences Cognitives (LNC) - CNRS - UMR7291 de Marseille, et en particulier dans le cadre d'un projet financé par l'Agence Nationale de la Recherche, le projet MetaSon : Métaphores Sonores. Un des objectifs de ce projet est le développement d'un stylo sonore permettant d'aider les enfants souffrant de dysgraphie, pathologie motrice ayant pour symptômes des difficultés dans l'apprentissage de l'écriture (Hamstra-Bletz et Blöte, 1993; Smits-Engelsman *et al.*, 2001). Les travaux présentés dans les sections précédentes ont donc été également réalisés dans le but de fournir des outils de sonification pour la remédiation de leur geste d'écriture. Dans cette section, nous présenterons la dysgraphie et les travaux en cours menés en collaboration avec LNC mettant en pratique les méthodes de sonification présentées dans la section précédente pour la réhabilitation de la dysgraphie.

5.3.1 La dysgraphie : une pathologie graphomotrice

La dysgraphie est une pathologie motrice impliquant des difficultés dans la production du mouvement d'écriture chez certains enfants. Ses origines sont encore méconnues, et la dysgraphie n'est pas toujours détectée facilement. En effet, si un enfant atteint de ce type de trouble met beaucoup plus de temps que les autres pour produire une lettre, la différence n'est pas nécessairement visible sur la production graphique finale. On peut voir sur la figure 5.11 le mot *lapin* écrit par un enfant dysgraphique et un enfant n'ayant aucun problème moteur (dit normo-scripteur). On constate que l'enfant dysgraphique arrive à écrire convenablement le mot, ce qui rend la détection de cette pathologie difficile. Il est donc nécessaire de prendre en compte les paramètres dynamiques de l'écriture et réhabiliter ce type de pathologie. La sonification de l'écriture a donc deux objectifs distincts, d'une



FIGURE 5.11 – Exemple d'écriture du mot *lapin* par un enfant sans problème moteur (à gauche) et par un dysgraphique (à droite).

part donner un outil pour détecter la dysgraphie en la sonifiant, et d'autre part créer un *stylo sonore* qui permette de guider les dysgraphiques afin qu'ils retrouvent un geste fluide.

5.3.2 Stratégie de sonification

Pour le second objectif qui nous intéresse plus particulièrement ici, nous souhaitons notamment améliorer la lenteur des du geste d'écriture des enfants dysgraphiques par le biais d'un retour sonore. Pour cela, la même stratégie de sonification que celle adoptée dans l'expérience précédente a été utilisée. Dans le cas particulier de la dysgraphie, on peut noter que le descripteur caractérisant la disflueance semble particulièrement adapté dans la mesure où l'étude qui l'a mis en évidence a été réalisée en comparant le nombre de pics surnuméraires chez des enfants normo-scripteurs et chez des dysgraphiques.

5.3.3 Premières pistes pour la réhabilitation de la dysgraphie par le biais d'un retour sonore

L'expérience précédente a donc mis en évidence l'intérêt du son pour l'apprentissage de nouveaux caractères dans une tâche de production graphique chez des sujets normo-scripteurs. Des expériences sont actuellement menées par le LNC avec des enfants dans l'objectif d'évaluer l'intérêt de tels outils de sonification par rapport aux méthodes déjà existantes. Des résultats préliminaires (Danna *et al.*, 2013b) ont mis en évidence que la sonification avait des effets bénéfiques pour la remédiation de la dysgraphie. Cependant, et pour des raisons éthiques évidentes, cette expérience n'a pas encore pu être réalisée avec un groupe contrôle d'enfants dysgraphiques apprenant des lettres avec des méthodes classiques. Des expérimentations sont actuellement en cours pour évaluer et comparer les effets d'un apprentissage avec un retour sonore et d'un apprentissage sans retour sonore.

5.4 Conclusion

Dans ce chapitre, nous avons donc présenté deux applications des résultats fondamentaux mis en évidence dans les trois chapitres précédents. La première application concerne la création de sons in-

ouïs, aussi appelés métaphores sonores, évoquant des situations a priori impossibles en se basant sur la flexibilité du paradigme perceptif action → objet. D'autre part, nous avons présenté une stratégie de sonification pour l'apprentissage du geste d'écriture basée sur le contrôle d'un modèle de synthèse de friction non linéaire. Cette stratégie a été évaluée dans le cadre d'une preuve de concept et est actuellement mise à l'épreuve dans le cadre de la réhabilitation d'une pathologie grapho-motrice : la dysgraphie. Des perspectives d'applications plus générales pour l'apprentissage d'un geste seront présentées dans la conclusion générale.

Chapitre 6

Conclusion générale & Perspectives

Sommaire

6.1	Vers d’autres relations entre le son et le mouvement biologique	140
6.2	Vers des stratégies de sonification pour le guidage et l’apprentissage de gestes . .	143
6.3	Vers un véritable paradigme d’étude des relations entre gestes et sons	145

Dans cette thèse, l’objectif était donc d’étudier les relations entre sons et gestes avec une approche très fondamentale. Pour cela, nous avons considéré le cas particulier des gestes graphiques. Ces gestes peuvent sembler silencieux mais pourtant, quand une personne dessine, des sons de frottement sont produits, et bien que l’on n’y prête que rarement attention, les différentes études que nous avons menées ont montré qu’ils contenaient en fait beaucoup d’informations pertinentes du point de vue perceptif pour reconnaître les caractéristiques, aussi bien cinématiques que géométriques, du geste. De façon très générale, nous avons pu mettre en évidence une relation très forte entre les sons et les mouvements biologiques qui était jusqu’alors insoupçonnée.

Pour aborder ce problème, nous avons utilisé une approche originale basée sur des outils de synthèse sonore s’appuyant sur un paradigme issu de l’approche écologique de la perception, le paradigme action → objet, qui stipule qu’un son du quotidien comme un son de frottement est le résultat de l’action d’un excitateur, par exemple la mine du stylo, sur un objet résonnant, la table par exemple. Pour contrôler ces modèles de synthèse, nous nous sommes essentiellement appuyés sur des modèles de geste mettant en évidence une invariance entre la cinématique d’un mouvement et la courbure d’une forme dessinée, la loi en puissance 1/3. Notre système perceptif, et plus particulièrement notre système auditif s’appuyant essentiellement sur ces informations invariantes pour reconnaître et identifier un événement sonore. Nous avons donc fait l’hypothèse qu’il était peut-être possible de reconnaître ces invariants gestuels à travers les sons de frottement et ainsi montrer une relation forte entre les gestes et les caractéristiques spectro-temporelles des sons.

Dans un premier temps, nous avons pu mettre en évidence que les invariances cinématiques, et plus particulièrement la loi en puissance 1/3 qui caractérise les mouvements biologiques, pouvaient être reconnues par le biais de l’évocation du profil de vitesse à travers un son de frottement. De plus, il a également été mis en évidence que des formes dessinées pouvaient être discriminées du point de vue sonore si elles sont assez distinctes du point de vue géométrique, et donc cinématique. Dans une seconde étude, nous nous sommes alors focalisés sur une forme particulière, l’ellipse, afin de comprendre les relations entre les sons de frottement et les caractéristiques géométriques réellement imaginées par le sujet. Pour cela nous nous sommes basés sur un modèle de geste considérant

le mouvement comme la conséquence du couplage entre deux oscillateurs harmoniques déphasés. Nous avons ainsi pu mettre en évidence l'existence d'un prototype dynamique sonore correspondant à la catégorie de forme ellipse. Et de plus, ce prototype dynamique correspond à la géométrie des prototypes elliptiques que nous avons également mis en évidence aussi bien du point de vue visuel que moteur. Cependant, nous avons également montré l'existence d'une distorsion entre l'ellipse, et donc le profil de vitesse utilisé pour synthétiser le son de frottement, et la forme, et donc le profil de vitesse imaginé par les sujets à partir du son. Ces résultats ont donc révélé une différence entre les métriques perceptives visuelles et sonores du point de vue de la perception de la cinématique. Cette différence entre perception visuelle et sonore de la dynamique a été évaluée dans une troisième étude lors de laquelle on a pu montrer que le couplage visuo-moteur avec un mouvement visuel était significativement affecté par l'ajout d'une dynamique sonore incongruente. Plus précisément, la géométrie de la forme visuelle reproduite en boucle ouverte visuelle a été significativement distordue par l'ajout d'une dynamique sonore. Ce résultat montre donc le rôle essentiel de la modalité auditive dans l'intégration multisensorielle d'un mouvement continu ce qui n'avait jamais été mis en évidence de cette façon à notre connaissance. Finalement, au regard de tous ces résultats fondamentaux sur les relations entre gestes et sons, et grâce au paradigme de synthèse basé sur l'existence d'invariants reliés à l'action et à l'objet, nous avons pu proposer des outils de créations de métaphores sonores, pouvant aller jusqu'à évoquer des situations a priori impossibles. De plus, une stratégie de sonification du geste d'écriture basée sur un modèle de synthèse de friction non linéaire a été proposée et évaluée dans le cadre d'une preuve de concept qui a pu montrer qu'apprendre des nouveaux caractères avec le son bénéficiait à l'apprentissage d'autres caractères sans le son.

Finalement, toutes ces études ouvrent des perspectives dans les différents domaines touchant à cette thèse. Ces perspectives vont être présentées dans ce dernier chapitre.

6.1 Vers d'autres relations entre le son et le mouvement biologique

Les approches sensori-motrices, et en particulier l'approche écologique, présentées en introduction nous ont donc permis d'aborder le problème de la relation entre le son et les gestes en formulant l'hypothèse forte que notre perception se base essentiellement sur la connaissance des caractéristiques du système moteur pour la formation des percepts à partir du son de frottement produit quand une personne dessine. Cependant, la relation entre sons et gestes pourrait être étendue à d'autres paramètres gestuels et éventuellement mettre en évidence d'autres invariants.

6.1.1 D'autres invariants associés aux gestes humains ?

Dans nos études nous nous sommes particulièrement focalisés sur la cinématique des mouvements biologiques. En comparant les résultats de tests perceptifs réalisés avec des sons de frottement enregistrés et des sons synthétisés à partir du profil de vitesse, nous avons pu montrer que le profil de vitesse des gestes est un invariant transformationnel de la perception de l'action réalisée sur un objet du point de vue sonore. Afin d'aller plus loin dans la compréhension des relations entre son et mouvements, il semble désormais intéressant de s'intéresser également à une autre caractéristique dynamique, que l'on pourrait appeler *pression* du geste. Physiquement, cette composante correspond à la force avec laquelle on appuie avec le stylo sur une feuille quand on dessine. Du point de vue du geste graphique, Schomaker et Plamondon (1990) ont montré qu'il n'existait pas de relation telle que la loi en puissance $1/3$ pour la pression, cette composante du geste semble beaucoup plus idiosyncrasique et il serait donc intéressant d'évaluer dans quelle mesure nous sommes capables de reconnaître cette grandeur dans des sons de synthèse générés à partir de l'enregistrement de la pression du sujet lui-même quand il dessine une forme. Du point de vue acoustique, la pression pourrait être reliée

au type d'interaction, on peut en effet imaginer que plus on appuie fort, plus la surface vue par la mine du stylo va être grande, et plus le son va évoquer un grattement. Ainsi, on pourrait relier les variations de pression à des transitions entre des sons de frottement et des sons de grattement. En pratique, on pourrait envisager des protocoles où l'on demande au sujet de reconnaître son geste parmi des sons de synthèse intégrant cette composante de force (dans des variations de timbre où d'intensité par exemple), les sons qui lui seraient proposés contiendraient des sons de son geste à lui mais également ceux d'autres sujets. Il pourrait ainsi être intéressant d'évaluer dans quelle mesure nous sommes capables de reconnaître des caractéristiques qui nous sont propres par rapport à celles d'autres individus. Kennel *et al.* (2014) ont en effet récemment été montré que des sportifs étaient capables de reconnaître les sons produits par leurs pas en franchissant une série de haies parmi des enregistrements contenant également ceux d'autres personnes parcourant la même série de haies.

Dans le chapitre 4 et en annexe 7.2, nous avons mené des expériences de synchronisation audio-motrices qui ont pu mettre en évidence que ce mapping n'avait d'influence ni sur la taille, ni sur la géométrie du geste évoqué par le son de frottement de synthèse. Cependant, de même que les études qui ont montré que la taille d'un objet peut être perçue à partir du son produit par impact (cf. section 1.2.1), il serait intéressant de déterminer dans notre cas s'il existe un invariant perceptif associé à l'évocation de la taille de la trace produite. Dans une expérience préliminaire présentée en annexe 7.4, nous avons pu montrer que si l'on demande explicitement à des sujets de modifier le mapping pour évoquer des formes plus ou moins grandes, alors ce paramètre pouvait évoquer des formes de tailles différentes. Ce résultat préliminaire suggère que les caractéristiques géométriques du geste évoqué par le son peuvent être différentes de la relation entre le son et le geste produit dans une tâche de synchronisation audio-motrice. Afin de déterminer plus précisément les invariants acoustiques reliés à la taille du geste, il pourrait être intéressant de réaliser d'autres expériences se focalisant sur l'influence précise du mapping sur le geste évoqué dans des tâches implicites ou explicites.

6.1.2 D'autres invariants associés aux mouvements humains ?

Dans cette thèse nous nous sommes focalisés sur les relations entre la perception sonore et les gestes graphiques. Les relations entre le mouvement d'une source sonore, les morphologies sonores, et les mouvements humains est encore un sujet très peu étudié qui reste à explorer et on peut imaginer que d'autres invariants perceptifs reliés à des mouvements du corps existent. On peut mentionner que des études menées actuellement entre le Laboratoire de Mécanique et d'Acoustique et l'Institut des Sciences du Mouvement s'intéressent à la relation entre la posture et le son. En particulier, dans une étude préliminaire, Gandemer *et al.* (2014) ont pu montrer que le mouvement d'un son spatialisé autour d'un sujet avait tendance à stabiliser sa posture (i.e. son centre de gravité postural). De plus, les oscillations posturales redevenaient plus grandes dès l'instant où il n'y avait plus de son. Cette étude montre donc que la relation entre le son et le mouvement est à considérer aussi bien au niveau des mouvements graphiques, mais de façon plus générale avec tous les types de mouvements que l'on peut réaliser. On peut par exemple penser aux mouvements oculaires ou encore aux mouvements de la tête. Le cas des mouvements oculaires semblent une piste très intéressante car comme les mouvements graphiques, le tracking visuel d'une cible en mouvement continu (se déplaçant en deçà d'une certaine vitesse) implique des mouvements qui respectent la loi en puissance 1/3 (de Sperati et Viviani, 1997).

6.1.3 Les émotions transmises par la musique : une simulation des mouvements biologiques ?

Dans cette thèse, nous avons donc montré que percevoir le son produit par une personne qui dessine, c'était s'imaginer ses caractéristiques cinématiques, et en particulier des caractéristiques biomécaniques comme la phase relative entre des oscillateurs harmoniques modélisant bien le fonctionnement des effecteurs moteurs. Dans les tâches que nous avons proposées, les sujets savaient explicitement qu'ils devaient tenter d'imaginer des caractéristiques motrices. D'une façon plus générale et dans un contexte artistique et émotionnel, il a été mis en évidence que les émotions transmises par une oeuvre d'art peuvent émaner d'une simulation motrice. Il semble ainsi intéressant de se demander si de façon implicite nous n'imaginons pas de telles caractéristiques quand nous entendons des sons abstraits tels que les morceaux de musique. Les émotions transmises par la musique pourraient ainsi être en partie reliées à une tentative de simulation de l'action par notre système cognitif.

Les émotions transmises par la musique sont un sujet d'étude très ancien et nous n'avons pas la prétention d'en faire une revue car ce n'est pas le sujet de cette thèse, nous souhaitons juste évoquer une hypothèse que nos résultats nous amènent naturellement à formuler. Si quand nous percevons les sons de frottement produits par une personne qui dessine nous sommes capables de nous imaginer le geste réalisé, on peut se demander si lorsque l'on écoute un morceau de musique nous n'accomplissons pas cette tâche implicitement ? Et de plus, si cette tentative de simulation de l'action ne pourrait pas être une dimension des émotions transmises par la musique.

*L'approche incarnée de la cognition musicale*¹ (Leman, 2008) propose en effet d'expliquer les émotions et interprétations impliquées dans l'écoute musicale par les notions d'*engagement* et de *résonance corporelle* qui correspondent à une description non linguistique des émotions transmises par la musique. Ces notions insistent sur le fait que notre corps est complètement impliqué dans la perception de la musique. Cette approche de la perception musicale dérive directement des approches sensori-motrices présentées en introduction de la thèse et en particulier de l'approche éactive proposée par Francisco Varela *et al.* (1993). Cela a également été suggéré dans d'autres domaines artistiques comme la peinture. Au regard des connaissances neuropsychologiques sur la simulation du geste quand nous percevons une forme, Longcamp et Velay (2010) ont par exemple proposé que les émotions que l'on a quand on voit une peinture pourraient émaner d'une tentative de simulation du geste de l'artiste qui a produit le tableau. Jeannerod (2011) a lui suggéré que les émotions transmises lorsque l'on assiste à un spectacle de danse peuvent émaner de la difficulté que l'on a à s'imaginer les mouvements qui sont en train d'être réalisés par le danseur, mouvements que nous sommes pour la plupart incapables de réaliser.

Au regard de nos résultats, on peut faire une hypothèse similaire pour la perception d'un morceau de musique. Une des origines possibles pourrait être une tentative de simulation d'un geste, et plus généralement d'un mouvement du corps, par rapport aux variations intrinsèques du timbre que l'on intégrerait comme des variations de vitesse. Cette hypothèse avait déjà été formulée dans un article de Shove et Repp (1995), à l'époque les travaux de Viviani sur la perception visuelle des mouvements biologiques étaient très récents et la relation avec la perception sonore n'était alors qu'une hypothèse. Des études récentes ont de plus mis en évidence le pouvoir des gestes pour décrire des extraits musicaux (Caramiaux, 2011). Nos résultats suggèrent donc clairement que cette relation entre la musique et les mouvements biologiques n'est peut-être pas du tout fortuite et qu'il convient de l'envisager plus sérieusement.

1. *Embodied music cognition* en anglais.

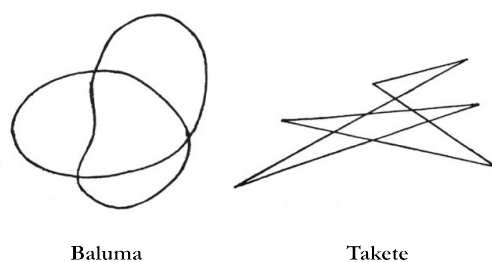


FIGURE 6.1 – Exemple de synesthésie entre des mots et des formes mis en évidence par Köhler (1929) – Si on demande à des sujets d'associer les deux formes ci-dessus aux mots *Baluma* et *Takete*, la quasi totalité des gens associe la forme de gauche au mot *Baluma* et celle de droite au mot *Takete*. Nous proposons ici que cette correspondance multimodale pourrait venir d'une correspondance entre les phases relatives émanant des gestes graphiques réalisés pour les dessiner et de celles des mouvements des gestes phonatoires réalisés pour prononcer les deux mots.

Concrètement, on peut faire l'hypothèse que percevoir une variation de barycentre spectral est perçue et intégrée en terme de phase relative entre deux oscillateurs harmoniques, correspondant à un mouvement et une trajectoire donnée. Les différentes qualités d'un mouvement, son aspect fluide ou saccadé par exemple², pourraient être reliées à des dimensions de l'émotion transmises par un extrait musical³. Cela permettrait de formaliser la notion de résonance corporelle dans une situation très précise. On ne prétend bien sûr pas expliquer toutes les émotions par le biais de cette relation entre mouvements biologiques et variations de barycentre spectral, les aspects rythmiques par exemple pouvant évoquer d'autres résonances corporelles (le rythme cardiaque par exemple).

Au-delà des émotions transmises par la musique, cette relation implicite entre son et geste pourrait également être envisagée dans une tentative d'explication de certaines synesthésies, par exemple l'association naturelle entre des formes et des mots (*Baluma* / *Takete* : voir figure 6.1). La théorie motrice de la perception suggérant que les mots sont perçus en terme de gestes phonatoires, l'association entre les phases relatives extraites des formes visuelles et celles impliquées dans la production vocale de ces mots⁴ pouvant ainsi expliquer ces associations qui semblent fortuites (Spence, 2011).

6.2 Vers des stratégies de sonification pour le guidage et l'apprentissage de gestes

Au regard des connaissances théoriques mises en évidences dans les chapitres 2, 3, 4, on a vu dans le chapitre 5 que les outils de synthèse de sons de friction linéaire (frottement, grattement...) et non linéaire permettaient d'envisager la création d'outils de guidage et d'apprentissage d'un geste. Une application concrète est actuellement testée dans le cadre d'une collaboration avec le Laboratoire de Neurosciences Cognitives de Marseille pour la réhabilitation de la dysgraphie. Cette application a été validée par une preuve de concept où des sujets devaient apprendre à dessiner une forme graphique qu'ils ne connaissaient pas avec leur main non dominante. D'une façon plus générale, ce paradigme de synthèse de sons de friction non linéaire peut permettre d'envisager la création d'outils de guidage des paramètres dynamiques d'un geste comme la pression et la vitesse.

2. Des expériences informelles réalisées durant la thèse mettent clairement en évidence qu'en fonction de la phase relative, la *qualité* du geste évoqué réfère très clairement à ce type de description sémantique.

3. On peut en particulier mentionner la description des émotions par le biais d'un diagramme *valence/arousal* (Bigand *et al.*, 2005).

4. Kelso *et al.* (1985); Kelso (1986) ont montré que les mouvements des lèvres, de langue, et de la mâchoire impliqués dans la production de la parole pouvant également être vus en terme de phase relative entre oscillateur.

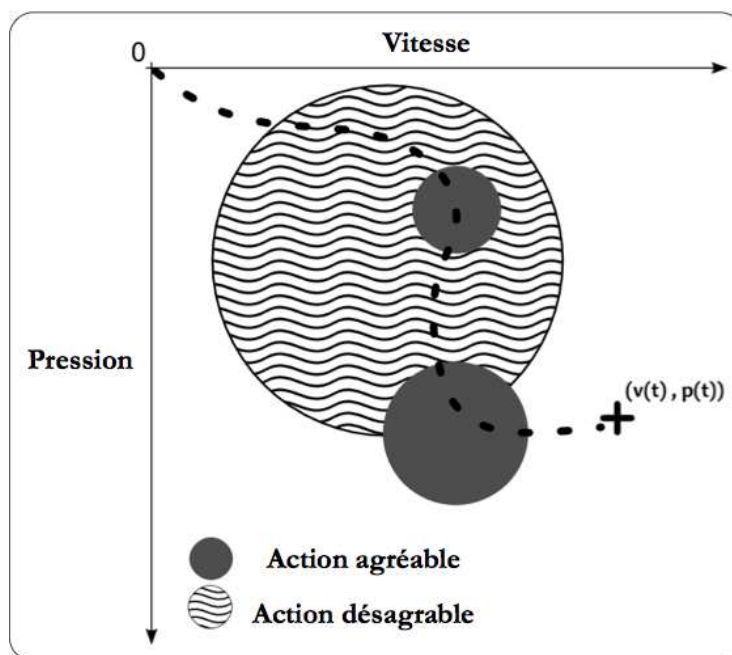


FIGURE 6.2 – Exemple de stratégie de contrôle entre les paramètres dynamiques $(v(t), p(t))$ et les différents états vibratoires : quand les paramètres du geste ne sont pas dans cible, un son *désagréable* (par exemple un grattement ou un couinement) est généré – quand les paramètres du gestes sont dans la cible, un sons *agréable* (par exemple un verre qui chante ou un frottement) est généré.

6.2.1 Vers le guidage du geste par le son

🔊 Le modèle de friction non linéaire s'est donc appuyé sur des considérations physiques et des observations acoustiques pour générer des sons de couinements, mais également des sons de verres qui chantent, ou encore de cordes frottées. Pour le moment, seul l'utilisation des sons de couinement a été envisagée dans le cadre du guidage, mais on pourrait également envisager l'utilisation des autres types de sons de friction non linéaire pour guider la vitesse et la pression de différents types de gestes. D'un point de vue très pragmatique et dans le cadre d'une preuve de concept très générale on pourrait envisager une tâche de guidage lors de laquelle un sujet devrait atteindre une cible de pression et de vitesse plus ou moins grande en dessinant continûment un cercle, qui pourrait correspondre de façon métaphorique au bord d'un verre, tout en ayant pour retour sonore soit un son de couinement si il est hors de la cible, soit un son de verre qui chante dès lors qu'il atteint la cible. Cela est schématisé sur la figure 6.2, on voit sur ce schéma qu'il est possible de définir un espace de guidage des paramètres pression / vitesse *a priori*. La flexibilité de cet espace de guidage permettant en outre de définir des zones plus ou moins grandes en fonction de l'expertise des sujets et de réduire la taille de ces zones au fur et à mesure que le sujet s'améliore.

Ce paradigme de contrôle est très proche de celui d'un instrument à cordes frottées comme un violon, en effet, quand on joue de ce type d'instrument, la justesse de la note jouée dépend de façon très forte des paramètres de pression et de vitesse de l'archet sur la corde. Et en fonction de ces paramètres physiques, les zones de pression et de vitesse où l'instrument produit un son correct sont plus ou moins grandes. On peut ainsi définir la jouabilité, i.e. on pourrait dire la facilité à produire un son correct, d'un instrument de musique en fonction de ces critères acoustiques (Young et Serafin, 2003). La jouabilité d'un instrument a été formalisée et est connue sous le nom de diagramme de Schelleng (1973). Ce type de diagramme représente le degré de jouabilité en fonction de la pression et de la vitesse de l'archet. Différentes études ont mis en évidence des zones où la jouabilité était

maximale et inversement. Notre espace de contrôle s'apparente donc à un diagramme de Schelleng qui pourrait être modifié à souhait pour créer des instruments virtuels plus faciles à jouer que des instruments réels. Dans la mesure où notre modèle permet de générer des sons de cordes frottées, il semble possible de le calibrer en fonction des paramètres acoustiques d'un vrai violon, et ainsi proposer un violon virtuel permettant un apprentissage avec différents niveaux de difficultés pour des musiciens non experts. Cette possibilité de créer des instruments virtuels dont on pourrait paramétrer la jouabilité a priori a déjà été envisagée et on peut en particulier mentionner une étude de Rován *et al.* (1997) sur la clarinette.

6.2.2 Sonifier d'autres invariants ?

Dans toutes les stratégies de sonification pour le guidage de nous avons soit testées soit évoquées jusqu'à présent, seuls des paramètres directement reliés au geste comme la vitesse ou la pression ont été considérés. Cependant, pour effectuer certaines actions spécifiques, notre système perceptivo-moteur semble s'appuyer également sur d'autres invariants perceptifs. Une théorie générale sur le guidage des mouvements a été émise par David Lee (1998), qui propose que pour guider une grandeur spécifique d'un mouvement (une force, une vitesse, un angle...), notre système perceptivo-moteur s'appuie sur le couplage entre le temps τ_r qu'il faut pour atteindre la grandeur cible et le temps interne τ_i que l'on s' imagine pour l'atteindre. Concrètement, le rapport $\frac{\tau_r}{\tau_i}$ est maintenu constant durant la totalité du mouvement. Ce modèle a été validé dans différentes situations, par exemple dans le cas du geste de put du golfeur. Craig *et al.* (2000) ont montré que chez des golfeurs professionnels, le temps nécessaire pour que l'extrémité du club atteigne la balle était égal à $\tau_r = k\tau_i = k\frac{t-T^2/t}{2}$ où t est le temps depuis le début de l'action et T la durée totale du put. Cette approche est très intéressante car elle tient compte à la fois de l'action imaginée et de l'action réellement produite, qui ont des caractéristiques communes comme on a pu le présenter dans l'état de l'art à travers les expériences d'imagerie motrices (Jeannerod, 1995). Tenir compte de telles contraintes perceptivo-motrices dans les tâches de sonification et de guidage audio-moteur s'avère donc une piste très intéressante et déjà mentionnée par Rodger et Craig (2011) dans des tâches de synchronisation audio-motrices.

6.3 Vers un véritable paradigme d'étude des relations entre gestes et sons

L'étude présentée dans le chapitre 4 a montré que le son pouvait fortement influencer le couplage visuo-moteur entre un mouvement visuel et une reproduction graphique. Cette expérience, mettant en jeu à la fois la vision, l'audition et la motricité permet d'envisager le développement d'un nouveau paradigme d'étude des relations entre gestes et sons. Pour cela, il serait nécessaire de consolider ces résultats par des expériences complémentaires. Il s'agit ainsi de formaliser le paradigme proposé et de démontrer que l'information auditive, en complément de l'information visuelle, module les boucles sensori-motrices à l'origine de la qualité du contrôle du geste.

6.3.1 Une nouvelle illusion audiovisuelle induite par l'audition

Dans le chapitre 4, nous avons pu mettre en évidence que la reproduction du mouvement d'un point lumineux en boucle ouverte visuelle (i.e. sans voir le mouvement que l'on reproduit) était très fortement impactée par l'ajout d'une dynamique sonore incohérente. Concrètement, quand un point parcourait une trajectoire visuelle circulaire à vitesse constante en présence d'un son évoquant des accélérations d'un mouvement biologique elliptique, les sujets reproduisaient alors une forme elliptique ne correspondant pas à la forme visuelle réellement perçue. De même, quand une ellipse était parcourue avec un mouvement biologique en présence d'un son évoquant une vitesse constante

d'un mouvement biologique circulaire, l'ellipse reproduite était alors plus circulaire que lorsque le son évoquait des accélérations synchrones avec les variations de vitesse visuelle. Cela montre que le couplage visuo-moteur est complètement relié à la modalité auditive et en particulier aux variations de timbre d'un son. Cette relation n'avait jusqu'alors jamais été mise en évidence aussi clairement à notre connaissance. Seules des études sur la synchronisation sensori-motrice avec des stimuli discrets ((Repp et Su, 2013) pour une revue) ont déjà mis en évidence la dominance du son sur la vision et en particulier que l'on était plus attiré par des rythmes sonores que visuels (Repp et Penel, 2004). Il pourrait désormais être intéressant de reconsidérer ce type d'étude en sonifiant également *les mouvements en l'air*. Une étude de Rodger et Craig (2011) a en effet mis en évidence que la planification du mouvement était différente dans une tâche de synchronisation audio-motrice avec des stimuli discrets ou continus.

Par ailleurs, pour valider complètement ce paradigme d'étude de la perception audiovisuelle du mouvement, il est essentiel d'évaluer si le visuel est impacté par le son dans une tâche purement perceptive n'impliquant que des stimuli visuels et sonores. Ainsi nous pourrions montrer qu'indépendamment de la reproduction, c'est la perception de la forme géométrique perçue qui est modifiée. Cette vérification pourrait être réalisée en reproduisant les études présentées dans le chapitre 4 mais dans une tâche purement psychophysique purement perceptive. Le résultat devrait nous apporter toutes les données comportementales permettant d'envisager de nombreuses investigations fondamentales, en particulier des études d'imagerie cérébrales qui permettraient de comprendre les processus cérébraux sous-jacents.

Du point de vue purement sonore, il pourrait être intéressant de déterminer les processus cérébraux impliqués dans la perception de sons de frottement monophoniques, en effet, bien qu'évoquant des mouvements ces sons ne mettent en jeu aucun indice spatiale. En particulier, il pourrait être intéressant de déterminer s'ils relèvent a priori d'un traitement dorsal ou ventral, ou des deux (il a en effet été mis en évidence que ces deux voies de traitement était *perméables* (Michel *et al.*, 2007; Norman, 2002)). De plus, l'existence d'une voie ventrale et d'une voie ventrale distincte au niveau auditif est encore débattue (Arnott *et al.*, 2004; Giordano *et al.*, 2012; Griffiths et Warren, 2004), ce type de stimuli pourraient ainsi permettre d'investiguer ce type de problème. Du point de vue multimodale, on peut envisager des perspectives similaires, et en particulier, si l'on peut mettre en évidence purement visuelle, comprendre l'origine de cette illusion au niveau cortical permettrait de mieux comprendre la perméabilité entre les traitements ventraux et dorsaux. Toutes ces questions pourraient être abordées dans des expériences réalisées en imagerie fonctionnelle.

6.3.2 La perception auditive de la dynamique : une inférence bayésienne ?

Dans le chapitre 3, nous avons également mis en évidence une distorsion entre la métrique visuelle et sonore dans la perception de la vitesse et de la représentation de la forme géométrique sous-jacente. Ce résultat pose la question de la façon dont nous fusionnons les informations émanant des deux modalités pour former des représentations cohérentes du monde. Il serait par exemple intéressant de savoir quel son serait le mieux adapté à un geste spécifique. Afin d'asseoir notre paradigme d'étude de la perception audiovisuelle du mouvement, il serait ainsi nécessaire de modéliser le poids de chaque modalité dans la formation de ces percepts. Des études mathématiques ont d'ores et déjà mis en évidence que le système visuel semble intégrer les grandeurs cinématiques dans un espace affine (Flash et Handzel, 2007) bien adapté pour représenter les gestes humains. D'autre part, nous avons pu mettre en évidence que du point de vue sonore nous intégrons la cinématique d'un mouvement sur la base de grandeurs euclidiennes, telle que la vitesse tangentielle. Enfin, les modèles

les plus récents d'intégration multimodale ont mis en évidence que pour former une représentation cohérente du monde, notre système perceptif choisit l'interprétation la plus probable en s'appuyant sur des inférences bayésiennes (Besson *et al.*, 2011; Weiss *et al.*, 2002). La distorsion mise en évidence suggère ici que le système visuel, et sa métrique affine sous-jacente, influence de façon significative la représentation de la cinématique du point de vue sonore. Pour complètement valider notre paradigme, mais également pour pouvoir distordre le son de façon à évoquer un mouvement spécifique dans une tâche de sonification, il semble nécessaire à ce stade de modéliser l'intégration audiovisuelle du mouvement en termes probabilistes, et en particulier avec un modèle bayésien.

6.3.3 Déterminer la contribution des interactions sensori-motrices dans l'émergence des prototypes géométriques

Par ailleurs, nous avons mis en évidence l'existence d'un prototype dynamique sonore correspondant en terme de géométrie au prototype visuel statique et dynamique (motrice) et au prototype moteur. Ce résultat est très intéressant car il porte un regard nouveau sur la notion de prototype géométrique. Dans la littérature et en particulier au regard des résultats de l'étude de Theurel *et al.* (2012), les prototypes émergent principalement d'une expérience visuelle. Nos résultats suggèrent qu'en plus de cette expérience visuelle, il est nécessaire de considérer le couplage sensori-moteur dans l'émergence des prototypes. En particulier, l'approche dynamique permet de faire l'hypothèse que ces prototypes correspondent à des attracteurs du mouvement, reliant ainsi la géométrie du prototype à la dynamique motrice correspondante. Une revue récente de Kelso (2012) sur la coordination dynamique entre les processus cérébraux et les processus moteurs suggère ainsi qu'il existe un couplage structurel très fort entre les contraintes dynamiques motrices et le fonctionnement des processus cognitifs sous-jacents, l'ensemble de ce réseau imbriqué fonctionnant à la façon d'un réseau auto-organisé. Nos résultats vont dans ce sens et montrent que les prototypes peuvent effectivement émerger en partie de contraintes sensori-motrices. Si la corrélation n'implique pas nécessairement une causalité, il semble nécessaire de considérer les contraintes biomécaniques dans l'émergence de ces formes spécifiques. De plus, la corrélation entre la géométrie (i.e. l'aplatissement) du prototype de l'ellipse et celle du rectangle mise en évidence par Kalénine *et al.* (2013) ouvre une perspective très intéressante pour envisager le rôle de la motricité dans l'émergence des prototypes. Pour cela il sera nécessaire de modéliser les dynamiques motrices impliquées dans la production de formes *discrètes* comme les rectangles en terme de transitions entre phase relatives.

Cette approche de la cognition en terme de réseau auto-organisé évoluant au grès des contraintes motrices et de l'environnement est très intéressante car elle s'intègre complètement dans une perspective éactive en plaçant le sujet au coeur du processus d'émergence du monde à travers ses contraintes sensori-motrices. L'utilisation du son comme outil d'investigation du rôle de la dynamique dans l'émergence des percepts cognitifs semble être une voie très intéressante et jusqu'à présent très peu explorée.

Postface

Dans cette thèse nous avons donc mis en évidence une relation évidente entre la perception sonore et les mouvements biologiques. Différentes expériences, purement perceptives ou impliquant également la vision et la motricité, ont montré que notre système auditif était intrinsèquement relié aux caractéristiques cinématiques de nos gestes. Tous ces résultats ouvrent des perspectives dans de nombreux domaines, aussi bien fondamentaux qu'appliqués discutés et présentés dans chaque chapitre et la conclusion générale précédente. Dans cette ouverture, nous souhaitons insister sur l'aspect pluridisciplinaire de ce travail, pluridisciplinarité qui a permis d'aborder ces problématiques avec les outils et méthodes appropriés, mais surtout de nourrir nos intuitions sur les aspects les plus pertinents à explorer.

Initiés par des problématiques liées à l'acoustique et le contrôle intuitif de la synthèse sonore, ces travaux nous ont naturellement amenés à aborder très tôt des problématiques relevant de la psychologie cognitive et des sciences du mouvement. La pluridisciplinarité de ces travaux n'a pas été évidente à prendre en main, *s'asseoir entre deux (voire trois) chaises* n'étant effectivement pas très confortable, au moins au début. En effet, l'acousticien froisse souvent le psychologue en lui parlant de représentations mentales des sons. Cependant, dès l'instant où il fait les efforts nécessaires pour comprendre le sens faible et fort du mot représentation, il découvre alors la subtilité de ces notions et surtout la quantité d'intuition que les approches sensori-motrices peuvent procurer pour envisager ses problèmes sous un regard un peu différent. Par ailleurs, l'acousticien a lui aussi à apporter au psychologue et aux sciences du mouvement. La synthèse sonore, et plus précisément les paradigmes de contrôles intuitifs fournissent en effet des outils très puissants pour manipuler et contrôler les sons sur la base de principes perceptifs tout en les alimentant par des modèles gestuels, permettant ainsi de créer des stimuli parfaitement adaptés à l'étude de la cognition et de l'action.

Finalement, les résultats de cette thèse mettent clairement en évidence toute la richesse de l'approche pluridisciplinaire (au sens fort !). Sans avoir la prétention de définir des contours différents aux domaines de recherche existants, le contrôle intuitif de la synthèse sonore offre un cadre d'étude propice à l'interaction entre des questions fondamentales sur la sensori-motricité et des problématiques plus appliquées portant sur le contrôle des modèles de synthèse. Il apparaît ainsi essentiel d'aborder les problématiques de la synthèse et du contrôle des sons avec les méthodes issues de la psychologie cognitive. Réciproquement, la psychologie cognitive de l'audition ne pourra qu'y gagner à aborder les questions qui lui sont propres avec des outils de synthèse adaptés, permettant de manipuler les propriétés intrinsèques des sons et ainsi proposer des protocoles interactifs pour évaluer le ressenti des sujets dans des situations *écologiques*. De même, en utilisant des paradigmes de synthèse intégrant le geste au sein même du processus de synthèse, les sciences du mouvement pourront créer de nouveaux paradigmes d'étude et d'apprentissage de gestes intégrant la modalité auditive, modalité encore souvent oubliée.

Chapitre 7

Annexes

7.1 Annexe 1 – Le mouvement de Lissajous respecte la loi en puissance 1/3

Dans cette annexe, nous allons tout d'abord établir la relation entre l'excentricité de l'ellipse et la phase relative entre les deux oscillateurs harmoniques couplés. Dans un second temps, nous montrons que, quelle que soit la phase relative, un mouvement de Lissajous composé de deux oscillateurs harmoniques couplés de même amplitude déphasés respecte la loi en puissance 1/3 quel que soit le déphasage et la fréquence.

Soit le mouvement de Lissajous suivant :

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} A \cos(\omega t) \\ A \cos(\omega t + \phi) \end{pmatrix}$$

où ϕ est compris entre 0 et 2π . Afin de faire coïncider le demi grand axe avec l'axe des abscisses, on effectue une rotation de $\frac{\pi}{4}$:

$$R(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$$

Après rotation des coordonnées $(x(t), y(t))$, on obtient finalement :

$$R\left(\frac{\pi}{4}\right) \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \frac{\sqrt{2}}{2} \begin{pmatrix} \cos(\omega t) - \cos(\omega t + \phi) \\ \cos(\omega t) + \cos(\omega t + \phi) \end{pmatrix} = \sqrt{2} \begin{pmatrix} -\sin(\frac{\phi}{2}) \sin(\omega t + \frac{\phi}{2}) \\ \cos(\frac{\phi}{2}) \cos(\omega t + \frac{\phi}{2}) \end{pmatrix}$$

Cette formulation nous permet donc de calculer l'excentricité e en fonction de la phase relative ϕ :

$$\boxed{e(\phi) = \sqrt{1 - \tan^2\left(\frac{\phi}{2}\right)}} \quad (7.1)$$

Nous allons maintenant montrer que ce type de mouvement respecte la loi en puissance 1/3 reliant la vitesse tangentielle v_t à la courbure C . La vitesse tangentielle v_t peut s'exprimer par

$$v_t(t) = \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}$$

par ailleurs l'expression de la courbure au cours du temps s'exprime par :

$$C(t) = \frac{|\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)|}{v_t(t)^3} \quad (7.2)$$

où \dot{x} , \dot{y} et \ddot{x} , \ddot{y} sont respectivement les dérivées première et seconde des coordonnées du point au cours du temps qui s'écrivent :

$$\begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix} = \begin{pmatrix} -A\omega \sin(\omega t) \\ -A\omega \sin(\omega t + \phi) \end{pmatrix}$$

et

$$\begin{pmatrix} \ddot{x}(t) \\ \ddot{y}(t) \end{pmatrix} = \begin{pmatrix} -A\omega^2 \cos(\omega t) \\ -A\omega^2 \cos(\omega t + \phi) \end{pmatrix}$$

on peut donc exprimer le numérateur de la courbure par :

$$|\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)| = A\omega^3 |\sin(\omega t) \cos(\omega t + \phi) - \cos(\omega t + \phi)| = A\omega^3 \sin(\phi)$$

en injectant cette expression dans l'équation 7.2, on obtient $C(t) = \frac{\omega^3 \sin(\phi)}{v_t(t)^3}$ soit finalement :

$$\boxed{v_t(t) = \omega \sin^{1/3}(\phi) C(t)^{-1/3}} \quad (7.3)$$

pour un ϕ et un ω donné, la grandeur $K(\phi, \omega) = \omega \sin^{1/3}(\phi)$ étant constante, on retrouve bien l'expression de la loi en puissance 1/3 quelle que soit ϕ : $v_t(t) = K(\phi, \omega) C(t)^{-1/3}$.

7.2 Annexe 2 – Effet du mapping sur la synchronisation audio-motrice

Cette annexe reprend en partie les résultats et analyses présentés dans le chapitre d'ouvrage et présente une expérience qui évalue l'influence du mapping entre la vitesse tangentielle et la fréquence de coupure du filtre passe-bas du modèle de synthèse de sons de frottement dans une tâche de synchronisation audio-motrice :

Thoret, E., Aramaki, M., Bourdin C., Bringoux L., Kronland-Martinet R., and Ystad S. (2014). Audio-motor synchronization : the effect of mapping between kinematics and acoustic cues on geometric motor features. *LNCS vol. 8905, Sound, Music and Motion, Springer Verlag Heidelberg*

Rappel sur le mapping

Cette section rappelle rapidement les principes de la synthèse des sons de frottement et en particulier comment ces sons ont été contrôlés, par une mise en correspondance entre les paramètres cinématiques et les paramètres du modèle de synthèse. Le modèle de Gaver (1993b); Van Den Doel *et al.* (2001) propose que les sons de frictions peuvent être synthétisés en considérant qu'ils sont la conséquence d'une suite d'impacts générés par l'interaction entre un excitateur et les aspérités d'une surface. Plus l'excitateur se déplace vite, plus le nombre d'impacts dans un intervalle de temps donné est élevé, et plus le pitch du son est élevé. Un tel modèle peut être implémenté en filtrant passe bas (ou passe bande) un bruit représentant la surface frottée¹. Conan *et al.* (2014b) ont récemment proposé une implémentation de ce modèle permettant de naviguer continûment entre différentes interactions (frotter-gratter-rouler). La modélisation du résonateur est réalisée en filtrant le bruit (déjà filtré par rapport au geste) par une banque de filtres résonnants dont les fréquences et amortissements sont ajustés en fonctions des caractéristiques modales de l'objet frotté (Aramaki et Kronland-Martinet, 2006; Aramaki *et al.*, 2010, 2011).

La simulation d'un mouvement par ce modèle de synthèse nécessite donc la détermination d'un mapping entre les paramètres dynamiques du geste, en particulier la vitesse tangentielle, et les paramètres du modèle de synthèse représentés par la fréquence de coupure du filtre passe-bas (ou de la fréquence centrale du filtre passe-bande) dans ce modèle. Cette fréquence f_c est reliée à la vitesse tangentielle v_t du mouvement par un coefficient de proportionnalité α : $f_c(t) = \alpha v_T(t)$. Ce coefficient rend compte de deux phénomènes physiques : d'une part la rugosité de la surface, mais également de la vitesse relative entre les deux objets en contact, ici entre la mine du stylo et les aspérités de la feuille de papier. La relation entre ces paramètres peut être illustrée par l'exemple suivant. Considérons une surface simple, avec des aspérités régulièrement espacées d'une distance d_0 . La figure 7.1 illustre que des variations fréquentielles du signal modélisé peuvent provenir de l'espacement entre les aspérités ou de la vitesse de frottement, ce qui peut induire une ambiguïté au niveau de l'interprétation sonore. En particulier, lorsqu'on double le rapport fréquence/vitesse, imagine-t-on une surface deux fois plus rugueuse ou une frottement deux fois plus rapide ?

Afin d'aborder ces questions, plusieurs expériences vont être présentées dans les sections qui suivent, la première portera sur la synchronisation entre un mouvement graphique et un son de frottement évoquant soit un geste elliptique soit un cercle. Le but sera de voir dans une telle tâche si le mapping a une influence sur les caractéristiques géométriques et cinématiques du geste.

1. Le modèle le plus classique pour modéliser la rugosité d'une surface est celui des bruits fractaux dont le spectre est défini par $S(\omega) = \frac{1}{\omega^\beta}$. Quand $\beta = 0$ le bruit est blanc, quand $\beta = 1$ le bruit est rose. Plus β est élevé, plus la surface est lisse.

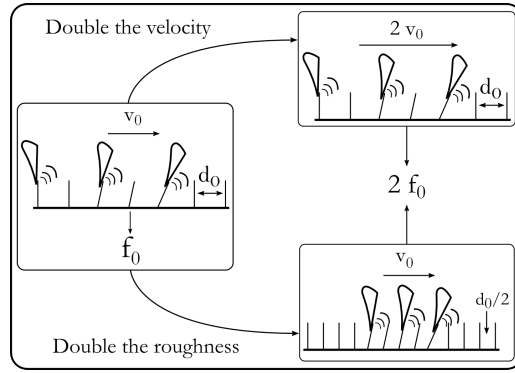


FIGURE 7.1 – Ambiguïté liée à une modification du mapping : multiplier le mapping par 2 peut soit être perçu comme un doublement de la rugosité de la surface frottée ou comme un doublement de la vitesse de l’excitateur.

Le but de l’expérience est donc d’évaluer l’influence du mapping sur la perception de la vitesse du mouvement et la rugosité de la surface dans une tâche de synchronisation audio-motrice en boucle ouverte visuelle, i.e. sans voir leurs mains. Cette ambiguïté entre vitesse et rugosité perçue pourra être évaluée à travers la taille des reproductions. En effet, si la taille des formes dessinées est influencée par le mapping, cela signifiera que la vitesse moyenne du mouvement évoquée est liée au mapping et donc que ce paramètre est intégré comme une variation de vitesse et non comme une variation de la rugosité de la surface.

Méthode

Sujets. 12 sujets ont pris part à cette expérience (2 femmes) d’âge moyen 24.17 (SD=2.55). Tous les participants étaient droitiers. Aucun d’entre eux n’était familier avec le sujet de l’expérience.

Appareillage. Les sujets avaient les yeux bandés et étaient assis devant un bureau. Les sons ont été diffusés avec un casque Sennheiser HD-650 et les gestes ont été enregistrés à l’aide d’une tablette graphique Wacom Intuos5 à un taux d’échantillonnage temporel de 133 Hz avec une précision spatiale de $5 \cdot 10^{-3}$ mm.

Stimuli. 

- *Formes géométriques.* Deux formes géométriques ont été utilisées. Une ellipse d’excentricité 0.9 de demi-grand axe 9.04 cm, et un cercle (i.e. ellipse d’excentricité nulle) de diamètre 12.72 cm (i.e. ellipse de demi-grand axe 3.18 cm). Les périmètres étaient de 43.86 cm pour l’ellipse et 40 cm pour le cercle.
- *Profils de vitesse.* La dynamique du mouvement le long de ces formes a été définie selon le mouvement de Lissajous suivant (i.e. il est important de noter que ce type de mouvement respecte la loi en puissance $1/3$, voir annexe 7.1) :

$$\begin{cases} x(t) = a \cos(\frac{2\pi}{T}t) \\ y(t) = b \sin(\frac{2\pi}{T}t) \end{cases} \quad (7.4)$$

où a et b sont respectivement le demi-grand et demi-petit axe des deux ellipses, $a = b$ dans le cas du cercle. La période choisie fait 1.8 s, et 19 périodes ont été générées. Les stimuli sonores ont ainsi une durée de 34.2 secondes. Le profil de vitesse de l’ellipse varie entre 13.88 cm.s^{-1} et 22.7 cm.s^{-1} , le profil de vitesse du cercle a une valeur constante de 22.7 cm.s^{-1} .

- *Stimuli acoustiques.* Les sons de frottement ont été synthétisés avec le modèle phénoménologique présenté précédemment, et à partir des profils de vitesse calculés à partir des mouvements définis dans le paragraphe précédent. 6 valeurs de α ont été choisies à cet effet : 5, 10, 20, 50, 100 et 300 m^{-1} . Ces valeurs permettent de générer des sons avec des variations de timbre (i.e. variations du barycentre spectral) vraiment différentes. La table 7.1 présente les valeurs minimales et maximales de la fréquence de coupure impliquées par les différentes valeurs de α . Finalement, 12 stimuli ont été générés : 2 (formes) \times 6 (mappings).

TABLE 7.1 – Valeur minimales et maximales de la fréquence de coupure du filtre passe-bas (en Hertz). Dans le cas du cercle la fréquence de coupure est constante durant toute la durée des stimuli.

	Ellipse		Cercle
α	f_{min}	f_{max}	f
5	69	159	111
10	138	318	223
20	276	635	445
50	690	1588	1135
100	1380	3176	2227
300	4140	9528	6681

Tâche et Procédure. La tâche consistait à dessiner une forme – un cercle ou une ellipse – en synchronisant son geste sur les variations de vitesse évoquées par le son de frottement diffusé dans le casque. Il était précisé aux sujets qu'ils devaient synchroniser leur geste sur les variations du son dans le sens anti-horaire pendant les 34.2 secondes du son. Il n'était pas précisé que les variations de timbre du son évoquaient les variations de vitesse du mouvement. Afin d'évaluer la relation directe entre la modalité auditive et le geste évoqué, les sujets avaient les yeux bandés tout au long de l'expérience. Ils étaient incités à transcrire les variations de timbre dans les caractéristiques de leur production. Il leur était explicitement demandé de lever leur coude de la tablette et de bloquer leur buste afin que leur mouvement soit généré uniquement par leur avant-bras. Les sujets pouvaient également ajuster la hauteur de leur siège avant l'expérience. Le test était précédé d'une phase d'entraînement durant laquelle les sujets étaient entraînés à synchroniser leur geste avec un exemple de son de frottement correspondant à un cercle ou une ellipse. Il était explicitement dit aux sujets que ces deux formes étaient contenues dans le test. Finalement, chaque sujet réalisait 36 essais : 2 (formes) \times 6 (mappings) \times {3 répétitions}. Les sujets étaient incités à faire une pause au milieu de l'expérience.

Analyse des données. Les données enregistrées correspondent aux coordonnées de la position du stylet sur la tablette au cours du temps. L'analyse des données s'est portée principalement sur les données géométriques et cinématiques caractérisant les performances motrices. Un pré-traitement des données a été réalisé afin d'éliminer les artefacts numériques liés au taux d'échantillonnage qui apparaissent en calculant les dérivées temporelles des coordonnées. Les méthodes utilisées pour calculer les différents descripteurs ont été en partie décrites par Viviani et Schneider (1991); Stucchi et Viviani (1993); Viviani *et al.* (1997).

- *Pré-traitement.* Les données ont été lissées en utilisant un filtre de Savitzky-Golay. De plus, comme les sujets avaient les yeux bandés, leurs mouvements n'étaient pas consistants spatia-

lement durant la totalité du mouvement et une déviation spatiale à très basse fréquence est apparue. Dans le but d'éliminer cette dérive spatiale, un filtrage passe-haut à 0.5 Hz a été réalisé sur les coordonnées enregistrées. Les enregistrements ont duré 34.2 s. Comme les sujets ont mis quelques périodes pour se synchroniser sur les variations du son et produire un geste régulier synchrone avec les variations évoquées par le son, en particulier pour les ellipses, les six premières périodes ont été éliminées de l'analyse, 12 périodes ont été utilisées (ce qui correspond à 21.6 secondes).

- *Caractéristiques géométriques et cinématiques.*

- *Excentricité* : L'excentricité d'une ellipse est définie par la formule suivante : $e = \sqrt{\frac{a^2 - b^2}{a^2}}$, où a et b sont les demi-grands et demi-petits axes de l'ellipse. Afin de déterminer l'excentricité de l'ellipse moyenne dessinée à chaque essai pour chaque sujet nous avons utilisé une méthode proposée par Viviani *et al.* (1997). Cette méthode propose de considérer que chaque enregistrement $(x(t), y(t))$ correspond à un nuage de point, où chaque point a un poids égal, dont on peut calculer le tenseur d'inertie. Il est bien connu en mécanique du point qu'un tenseur d'inertie d'ordre 2 peut être modélisé par une ellipse dont les caractéristiques sont reliées aux valeurs propres et vecteurs propres du tenseur d'inertie diagonalisé. La méthode précise pour calculer ce tenseur d'inertie est détaillée dans le livre de Goldstein (1962).
- *Tilt* : Le tilt correspond à l'angle d'inclinaison de l'ellipse dessinée par rapport à l'axe horizontal et est calculé à partir des vecteurs propres du tenseur d'inertie. De même que pour l'excentricité, le tilt est calculé en moyennant les valeurs sur les 12 périodes.
- *Périmètre moyen* : Le périmètre moyen a été calculé en utilisant les 12 périodes sélectionnées.
- *Vitesse moyenne* : La vitesse moyenne a été calculée à partir des 12 périodes sélectionnées.

Résultats

Tous les résultats sont résumés dans la table 7.2. Pour chaque descripteur et chaque forme géométrique, une ANOVA à mesures répétées a été réalisée afin d'évaluer comment le mapping a influé sur la géométrie et la cinématique de la forme dessinée. Chaque effet significatif a été analysé avec un test post-hoc de Newman-Keuls.

Excentricité. L'excentricité de l'ellipse est significativement affectée par le mapping pour les cercles mais pas pour les ellipses, bien que l'effet soit quasiment significatif pour ce dernier ($F(5, 55) = 5.907, p < .001$ et $F(5, 55) = 2.295, p = .06$ respectivement). Dans le cas des cercles, les tests post-hoc révèlent que le mapping $\alpha = 5$ a impliqué des productions graphiques significativement plus plates que dans les 5 autres conditions ($p < .05$ pour toutes les comparaisons). De plus, le mapping $\alpha = 20$ a significativement plus aplati les cercles que la condition $\alpha = 300$, voir la figure 7.2.

Tilt. L'orientation n'a pas été significativement affectée pour les cercles et les ellipses ($F(5, 55) = .854, p = .52$ et $F(5, 55) = 1.695, p = .151$ respectivement), voir figure 7.3. On peut constater que l'orientation moyenne semble plus élevée pour les cercles que pour les ellipses. Un cercle n'ayant pas d'orientation, on parle ici de l'orientation de la forme reproduite qui n'est pas parfaitement circulaire, i.e. elliptique.

TABLE 7.2 – Moyennes marginales et erreurs-standards pour les trois descripteurs géométriques.

	α	\bar{e}	$\bar{P}rimtre$ (cm)	$\bar{T}ilt$ (degrés)	\bar{v} (cm/s)
Cercle	5	.72 \pm .044	7.4 \pm 1.11	55.4 \pm 7.23	4.26 \pm 0.991
	10	.63 \pm .030	7.0 \pm .92	47.9 \pm 5.78	4.06 \pm 0.817
	20	.63 \pm .047	7.4 \pm 1.03	50.7 \pm 8.46	4.36 \pm 1.013
	50	.63 \pm .042	6.9 \pm 1.03	38.7 \pm 6.69	4.34 \pm 1.178
	100	.59 \pm .029	7.6 \pm 1.10	47.6 \pm 6.28	4.5 \pm 1.067
	300	.55 \pm .032	5.3 \pm 1.23	48.9 \pm 7.08	3.42 \pm 1.324
Ellipse	5	.93 \pm .016	15.7 \pm 1.82	37.9 \pm 6.20	8.74 \pm 1.981
	10	.93 \pm .011	15.8 \pm 1.37	32.9 \pm 5.20	8.63 \pm 1.487
	20	.94 \pm .009	15.9 \pm 1.57	39.4 \pm 6.27	8.84 \pm 1.693
	50	.94 \pm .013	16.1 \pm 1.62	32.7 \pm 7.28	9.00 \pm 1.773
	100	.95 \pm .010	15.9 \pm 1.45	24.9 \pm 5.17	8.87 \pm 1.592
	300	.94 \pm .013	16.4 \pm 1.89	35.2 \pm 7.77	9.19 \pm 2.102

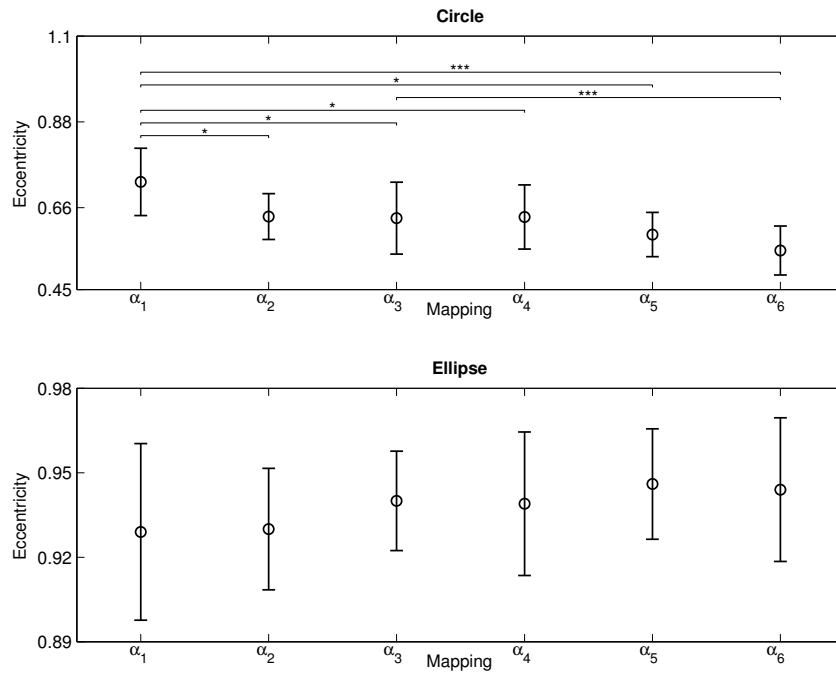


FIGURE 7.2 – Excentricités : moyennes marginales et intervalles de confiance à 95%

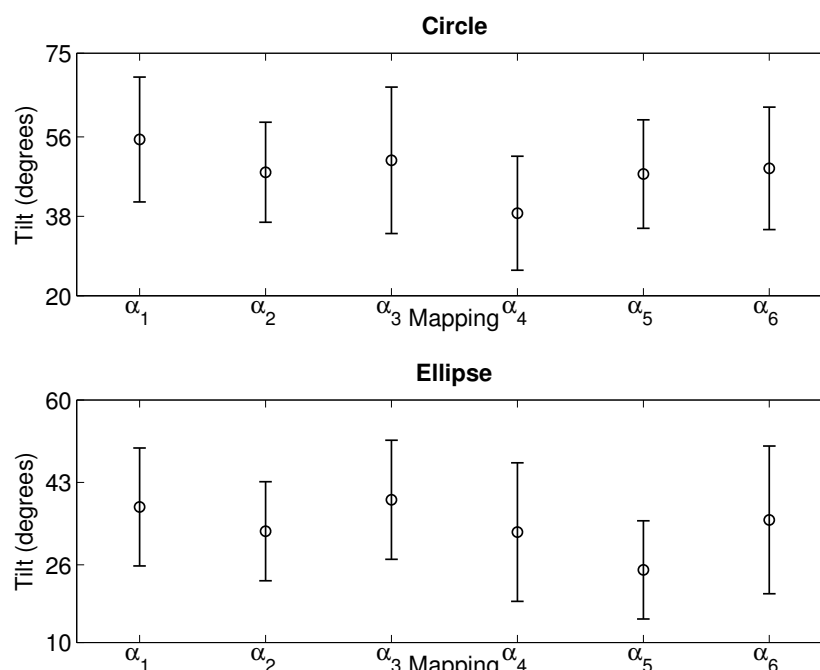


FIGURE 7.3 – Tilts (en degrés) : moyennes marginales et intervalles de confiance à 95%

Périmètre. Comme pour le tilt, aucun effet du mapping n'a été mis en évidence pour le périmètre ni pour les cercles ni pour les ellipses, ($F(5,55) = 1.642, p = .16$ et $F(5,55) = .569, p = .72$ respectivement), voir figure 7.4. Ces résultats montrent que, quel que soit le mapping, la taille de l'ellipse dessinée n'est pas significativement affectée, ce qui est vraiment intéressant car cela montre que dans cette tâche de synchronisation, le mapping n'est pas perçu comme une variation de vitesse mais comme un changement de rugosité, i.e. plus le mapping est élevé plus la rugosité est importante.

Vitesse moyenne. Aucun effet significatif du mapping n'a été mis en évidence ni pour les cercles ni pour les ellipses ($F(5,55) = 1.604, p = .17$ et $F(5,55) = .581, p = .71$ respectivement), voir figure 7.5.

Discussion

Dans cette expérience, les sujets devaient synchroniser leurs gestes sur des sons évoquant des gestes humains elliptiques ou circulaires. Le but était d'évaluer l'influence du mapping entre la vitesse tangentielle du mouvement et la fréquence de coupure du filtre passe-bas afin de déterminer si le mapping était d'avantage relié à la vitesse du geste ou à la rugosité de l'objet frotté. Les résultats ont ainsi montré que le mapping n'avait pas d'influence sur la taille de la forme reproduite. Cela montre que les sujets n'ont pas interprété les différents mappings comme différentes vitesses, sinon la taille de la forme aurait été modifiée.

Cependant, il est nécessaire de nuancer ce résultat, en effet on peut imaginer que la taille de la forme dessinée a été également contrainte par plusieurs autres facteurs. D'une part, la taille de la tablette a pu contraindre la taille de forme dessinée, et comme les sujets avaient les yeux bandés, il est possible qu'ils n'aient pas traduit complètement les variations de vitesse qu'ils ont pu extraire du son. D'autre part, on peut imaginer que les contraintes biomécaniques, en particulier la taille des membres impliqués dans le mouvement, ont également contraint la taille de l'ellipse dessinée. Le fait d'être en boucle ouverte visuelle a donc pu contraindre la taille de l'ellipse dessinée par rapport à ces deux considérations techniques et biomécaniques.

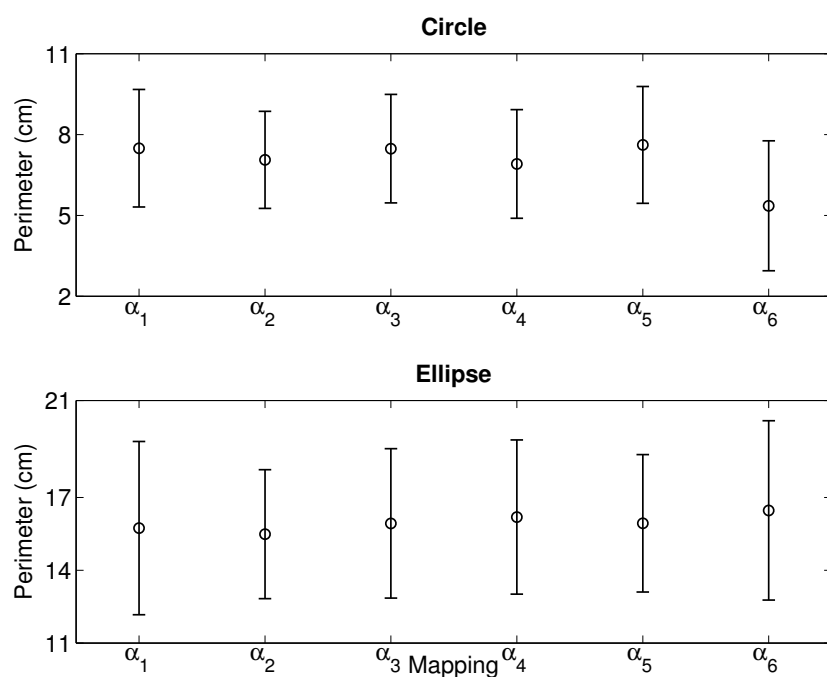


FIGURE 7.4 – Périmètres (en cm) : moyennes marginales et intervalles de confiance à 95%

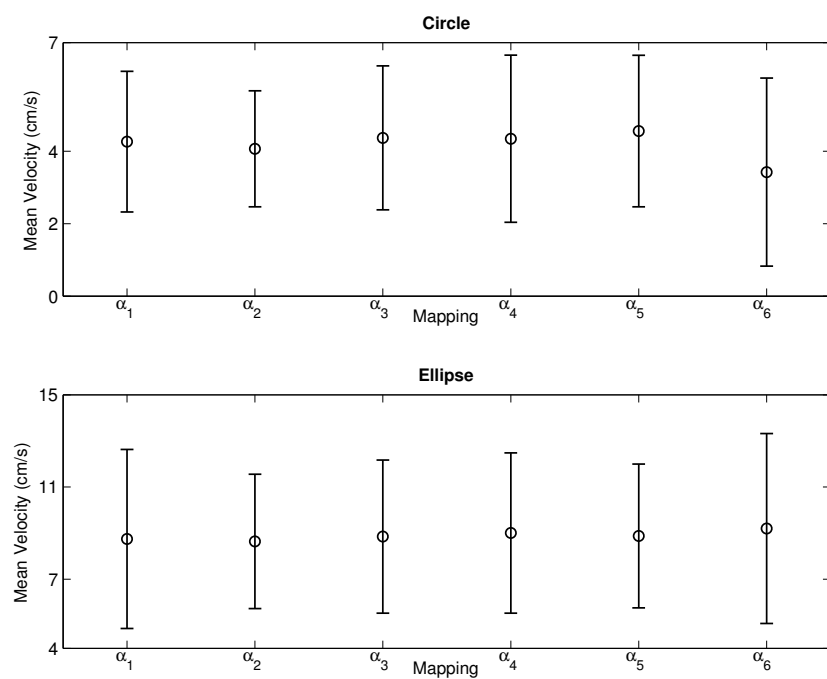


FIGURE 7.5 – Vitesses moyennes (cm/s) : moyennes marginales et intervalles de confiance à 95%

Il est également intéressant de constater que l'excentricité de l'ellipse dessinée est d'en moyenne .94 ce qui est une valeur légèrement supérieure à l'excentricité prototypique de .91 mise en évidence dans le chapitre 3. Cela peut être mis en regard avec les résultats du chapitre 3 qui ont montré que l'ellipse évoquée par les sons de frottement avait une excentricité plus élevée que celle réellement transmise dans le son. Si on applique la formule 3.10² correspondant à la distorsion entre l'ellipse sonore et visuelle, on trouve que l'ellipse sonore d'excentricité .9 correspond à une ellipse visuelle d'excentricité .96 qui est donc plus aplatie que l'ellipse initialement transmise dans le son. Les résultats de cette expérience sont donc cohérents avec la distorsion observée entre métrique visuelle et sonore dans le chapitre 3. Ce résultat est essentiel pour des applications comme la sonification de forme géométrique car il montre qu'il est nécessaire d'adapter le mapping pour évoquer la forme sonifiée.

2. Cette formule relie la phase relative ϕ_v de l'ellipse visuelle à la phase relative ϕ_s de l'ellipse sonore correspondante : $\phi_v(\phi_s) = 6.3721 - 0.2071\phi_s + 0.0186\phi_s^2 - 0.0001\phi_s^3$, cette formule peut être vue du point de vue de l'excentricité en utilisant la formule suivante $e = \sqrt{1 - \tan^2(\frac{\phi}{2})}$.

7.3 Annexe 3 – Résultats détaillés du chapitre 3

Cette section présente les résultats détaillés par sujet des expériences 4, 5, 6 et 7 du chapitre 3.

Expérience 4 – Prototype visuel statique

Sujet	A_1	A_2	A_3
S1	50.9 _(9.09)	45.9 _(6.08)	38.2 _(6.77)
S2	42.4 _(5.62)	49.0 _(7.98)	47.5 _(8.45)
S3	52.2 _(4.80)	43.4 _(5.61)	45.4 _(8.59)
S4	47.2 _(3.84)	51.4 _(4.62)	46.7 _(9.35)
S5	49.5 _(8.49)	42.5 _(5.46)	49.9 _(9.09)
S6	43.2 _(7.82)	51.4 _(4.64)	41.1 _(1.90)
S7	47.7 _(7.32)	37.6 _(5.40)	50.0 _(9.03)
S8	47.0 _(7.03)	51.7 _(6.26)	39.4 _(5.56)
S9	44.3 _(5.73)	41.9 _(5.08)	45.7 _(7.31)
S10	48.3 _(6.68)	48.2 _(7.77)	41.4 _(5.61)
S11	40.1 _(4.88)	41.5 _(6.48)	44.3 _(8.70)
S12	47.3 _(7.74)	45.3 _(6.71)	42.8 _(7.17)
S13	38.2 _(4.75)	44.1 _(7.95)	42.9 _(7.10)
S14	48.6 _(5.38)	41.0 _(7.96)	45.3 _(7.73)
S15	40.4 _(6.64)	47.9 _(7.09)	41.2 _(4.79)
S16	47.5 _(7.51)	35.9 _(8.07)	45.7 _(9.32)
S17	45.7 _(6.35)	53.2 _(6.83)	43.1 _(4.86)
S18	44.2 _(6.76)	40.2 _(8.80)	45.3 _(4.75)
S19	49.5 _(8.47)	50.4 _(7.61)	44.1 _(5.90)
S20	42.6 _(5.36)	45.3 _(6.72)	45.7 _(7.33)
Population	45.7 _(7.28)	45.5 _(7.52)	44.0 _(7.03)

TABLE 7.3: Résultats de l'expérience 4 – Prototype visuel statique : Médianes et JND (entre parenthèse) par sujets.

Expérience 5 – Prototype moteur

Sujet	A_1	A_2	A_3
S1	50.5 _(2.51)	54.1 _(2.89)	53.6 _(3.41)
S2	28.1 _(4.86)	27.2 _(3.78)	28.8 _(4.25)
S3	48.0 _(3.99)	48.3 _(3.69)	56.1 _(4.86)
S4	58.0 _(3.00)	51.8 _(3.20)	42.1 _(3.90)
S5	61.3 _(4.74)	58.1 _(3.94)	50.2 _(2.83)
S6	48.2 _(3.79)	52.1 _(3.92)	46.9 _(5.14)
S7	34.0 _(3.98)	36.5 _(4.47)	36.2 _(5.77)
S8	43.3 _(2.68)	45.2 _(3.81)	52.1 _(4.86)
S9	13.9 _(4.05)	21.7 _(3.29)	16.0 _(4.99)
S10	32.8 _(3.16)	37.5 _(3.50)	31.3 _(3.74)
S11	53.1 _(2.95)	46.1 _(2.89)	44.6 _(4.42)
S12	39.1 _(3.86)	34.3 _(3.70)	39.6 _(5.38)
S13	63.6 _(2.35)	61.1 _(2.87)	61.8 _(4.23)
S14	54.6 _(2.36)	50.5 _(3.47)	48.5 _(4.46)
S15	60.2 _(3.77)	47.3 _(3.68)	49.3 _(4.71)
S16	30.1 _(3.88)	27.0 _(4.03)	30.4 _(4.56)
S17	61.9 _(5.12)	58.2 _(2.84)	51.2 _(3.82)
S18	45.6 _(3.39)	42.9 _(3.13)	41.4 _(3.63)
S19	37.4 _(2.58)	40.1 _(2.91)	32.7 _(2.32)
S20	43.4 _(3.59)	58.4 _(4.61)	52.9 _(4.12)
Population	45.6 _(10.36)	45.6 _(8.3)	43.8 _(9.17)

TABLE 7.4: Résultats de l'expérience 5 – Prototype moteur : Médianes et JND (entre parenthèse) par sujets.

Expérience 6 – Prototype visuel dynamique

Taille	A_1			A_2			A_3		
Fréquence	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
S1	49.7 _(14.306)	51.6 _(16.406)	48.5 _(10.49)	45.3 _(10.67)	43.5 _(16.47)	49.2 _(11.77)	48.2 _(14.84)	49.6 _(13.35)	46.4 _(13.61)
S2	49.4 _(12.605)	53.9 _(16.120)	54.3 _(10.69)	48.5 _(14.50)	45.6 _(13.37)	43.4 _(13.63)	44.3 _(12.70)	43.7 _(17.30)	50.5 _(14.52)
S3	55.8 _(16.151)	47.0 _(8.995)	45.0 _(13.97)	46.5 _(11.53)	46.1 _(14.87)	48.5 _(14.52)	51.8 _(12.18)	44.4 _(11.57)	47.4 _(12.65)
S4	44.1 _(8.890)	49.7 _(12.261)	44.8 _(8.21)	52.0 _(12.03)	40.1 _(11.95)	48.0 _(17.02)	46.6 _(16.45)	48.1 _(14.91)	49.3 _(9.68)
S5	44.1 _(7.882)	44.0 _(14.048)	50.2 _(11.82)	47.5 _(13.48)	46.2 _(8.80)	33.1 _(14.90)	43.3 _(8.70)	52.1 _(11.92)	48.7 _(13.32)
S6	46.0 _(12.951)	39.1 _(10.874)	39.9 _(11.14)	45.9 _(12.05)	48.3 _(13.67)	46.3 _(11.69)	52.2 _(13.81)	45.9 _(8.09)	45.4 _(11.57)
S7	42.9 _(8.093)	44.2 _(12.777)	44.1 _(8.88)	49.1 _(6.89)	40.9 _(11.09)	44.4 _(12.58)	49.7 _(13.26)	48.8 _(9.17)	45.2 _(8.84)
S8	48.5 _(12.466)	40.9 _(12.074)	43.9 _(16.10)	44.4 _(12.64)	46.8 _(15.18)	44.1 _(13.89)	48.8 _(13.22)	48.0 _(13.04)	50.3 _(14.67)
S9	45.6 _(14.430)	47.6 _(10.384)	50.6 _(13.41)	45.9 _(13.11)	45.2 _(16.81)	47.3 _(12.70)	48.4 _(13.60)	46.2 _(11.83)	43.8 _(14.16)
S10	47.4 _(15.615)	47.2 _(13.805)	48.8 _(14.18)	53.4 _(13.65)	40.7 _(10.28)	50.7 _(13.31)	40.2 _(12.80)	51.1 _(15.91)	42.6 _(9.39)
S11	49.6 _(12.356)	42.9 _(11.146)	46.7 _(14.35)	48.3 _(16.68)	49.4 _(15.60)	49.2 _(6.84)	45.3 _(10.67)	42.6 _(13.43)	47.7 _(9.32)
S12	44.5 _(15.498)	50.8 _(11.188)	42.4 _(17.60)	48.5 _(11.52)	47.6 _(14.44)	49.9 _(12.10)	42.1 _(12.85)	41.5 _(9.49)	39.0 _(10.00)
S13	40.9 _(9.143)	44.2 _(17.792)	45.3 _(9.65)	45.2 _(10.78)	44.1 _(8.88)	43.6 _(12.37)	42.5 _(9.47)	43.2 _(13.76)	48.1 _(11.94)
S14	48.8 _(14.206)	42.5 _(9.544)	44.6 _(15.44)	43.6 _(11.44)	48.3 _(13.70)	41.2 _(9.80)	51.0 _(10.99)	45.5 _(14.50)	47.7 _(13.30)
S15	39.6 _(14.371)	45.5 _(14.459)	36.7 _(6.31)	38.3 _(7.75)	41.9 _(14.11)	43.7 _(11.26)	44.8 _(13.19)	40.2 _(9.85)	32.4 _(10.63)
S16	38.2 _(9.783)	42.4 _(14.561)	53.6 _(16.40)	45.3 _(12.70)	41.0 _(10.98)	40.8 _(11.25)	33.8 _(10.19)	42.8 _(14.24)	46.1 _(11.91)
S17	46.4 _(11.593)	36.4 _(10.584)	42.6 _(15.43)	38.1 _(9.91)	49.6 _(13.39)	33.6 _(9.40)	43.2 _(11.82)	33.4 _(9.63)	43.5 _(14.55)
S18	45.2 _(14.800)	46.7 _(10.324)	39.2 _(10.80)	36.3 _(7.72)	41.6 _(11.38)	43.9 _(15.09)	43.5 _(14.47)	41.7 _(9.32)	35.8 _(10.21)
S19	43.7 _(11.327)	44.7 _(15.276)	52.2 _(14.79)	47.3 _(11.67)	39.0 _(9.01)	34.9 _(13.11)	29.8 _(9.23)	45.1 _(13.94)	42.5 _(13.54)
S20	45.3 _(11.736)	39.8 _(9.176)	37.3 _(15.71)	37.2 _(9.83)	42.4 _(13.58)	40.1 _(12.91)	47.7 _(14.30)	33.2 _(10.82)	45.0 _(13.97)
Population	45.6 _(12.41)	44.8 _(13.18)	45.3 _(12.6)	45.1 _(11.92)	44.2 _(12.82)	43.8 _(13.20)	44.68 _(13.31)	44.19 _(12.81)	44.8 _(12.18)

TABLE 7.5: Résultats de l'expérience 6 – Prototype visuel dynamique : Médianes et JND (entre parenthèse) par sujets.

Expérience 7 – Prototype sonore

Taille	A_1			A_2			A_2		
Fréquence	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
S1	62.7 _(9.281)	52.4 _(13.592)	58.0 _(7.01)	55.3 _(15.73)	56.1 _(9.86)	64.5 _(8.48)	59.2 _(12.80)	58.4 _(11.62)	55.3 _(11.73)
S2	51.7 _(17.302)	54.5 _(11.533)	59.6 _(10.40)	63.4 _(11.62)	61.0 _(14.05)	54.2 _(6.83)	55.4 _(12.61)	61.2 _(8.82)	58.0 _(10.99)
S3	58.1 _(8.880)	57.9 _(10.144)	60.5 _(4.50)	55.3 _(13.73)	62.7 _(7.34)	55.5 _(9.46)	57.5 _(6.48)	55.3 _(15.75)	53.5 _(10.52)
S4	59.0 _(10.954)	58.7 _(13.334)	56.1 _(14.94)	57.5 _(14.48)	52.7 _(10.29)	51.9 _(15.15)	57.2 _(15.78)	54.0 _(12.01)	51.8 _(12.24)
S5	59.1 _(9.945)	59.0 _(8.017)	63.7 _(8.31)	56.0 _(11.00)	51.7 _(9.28)	59.2 _(9.80)	56.4 _(9.56)	61.5 _(9.55)	59.1 _(7.93)
S6	55.3 _(9.693)	59.5 _(7.546)	55.5 _(11.47)	53.2 _(13.76)	56.5 _(7.49)	53.6 _(8.42)	54.5 _(14.47)	54.3 _(8.70)	57.1 _(8.93)
S7	54.5 _(14.519)	54.5 _(8.476)	51.3 _(11.74)	34.5 _(17.54)	52.3 _(10.71)	54.6 _(7.36)	58.7 _(9.29)	61.5 _(11.47)	56.2 _(9.79)
S8	56.3 _(11.672)	59.9 _(10.108)	65.0 _(9.02)	65.2 _(9.85)	58.2 _(10.84)	60.8 _(10.20)	53.8 _(8.23)	59.9 _(8.10)	58.8 _(7.18)
S9	63.6 _(11.449)	59.9 _(16.075)	60.3 _(9.72)	60.9 _(11.09)	58.7 _(9.34)	59.7 _(9.26)	58.0 _(12.98)	56.7 _(10.29)	60.9 _(10.14)
S10	59.3 _(13.675)	66.8 _(7.235)	61.6 _(10.43)	62.9 _(10.13)	59.3 _(11.67)	63.9 _(7.15)	56.8 _(10.17)	54.6 _(11.42)	49.8 _(8.21)
S11	59.2 _(12.846)	61.1 _(6.869)	62.7 _(12.32)	63.1 _(8.86)	64.1 _(12.95)	54.9 _(11.11)	57.1 _(13.95)	57.5 _(9.52)	65.7 _(11.31)
S12	64.1 _(9.923)	63.9 _(13.135)	66.8 _(6.21)	55.3 _(9.74)	61.4 _(8.65)	58.9 _(9.13)	57.8 _(12.21)	60.4 _(7.60)	58.4 _(8.60)
S13	56.7 _(13.292)	56.1 _(7.859)	59.4 _(10.63)	58.6 _(10.41)	53.7 _(8.34)	55.3 _(13.71)	58.0 _(6.98)	48.8 _(15.22)	47.4 _(7.65)
S14	48.1 _(11.909)	52.7 _(8.327)	53.0 _(6.99)	63.1 _(8.87)	65.6 _(4.41)	57.0 _(15.04)	57.0 _(15.04)	56.0 _(7.00)	62.1 _(11.86)
S15	60.2 _(6.754)	58.8 _(15.214)	60.3 _(10.72)	59.9 _(12.10)	64.5 _(10.52)	63.9 _(9.09)	70.4 _(7.57)	60.7 _(12.32)	55.0 _(11.00)
S16	58.2 _(16.789)	60.6 _(8.379)	62.2 _(11.82)	64.0 _(6.96)	59.9 _(14.14)	61.8 _(9.18)	60.0 _(10.05)	64.7 _(12.34)	60.9 _(11.11)
S17	62.2 _(8.845)	61.3 _(9.723)	62.2 _(10.83)	59.2 _(8.83)	68.6 _(9.38)	58.5 _(11.50)	55.6 _(15.38)	61.4 _(9.62)	61.0 _(11.96)
S18	61.2 _(8.779)	64.7 _(11.320)	57.9 _(13.05)	57.3 _(14.66)	51.4 _(12.57)	61.7 _(11.28)	69.9 _(7.11)	63.4 _(12.63)	63.8 _(9.21)
S19	54.3 _(11.665)	59.4 _(9.627)	62.5 _(8.46)	62.5 _(8.54)	58.5 _(10.53)	57.2 _(9.75)	58.8 _(11.23)	62.3 _(12.74)	56.1 _(9.88)
S20	58.1 _(10.888)	52.5 _(8.470)	55.0 _(11.01)	61.9 _(9.08)	52.9 _(13.12)	55.2 _(10.76)	43.5 _(16.51)	52.5 _(11.47)	49.6 _(13.45)
Population	58.4 _(11.574)	58.6 _(10.370)	59.6 _(10.40)	59.0 _(11.97)	58.5 _(10.54)	58.3 _(9.73)	58.2 _(11.78)	58.3 _(11.71)	57.0 _(10.99)

TABLE 7.6: Résultats de l'expérience 7 – Prototype sonore : Médianes et JND (entre parenthèse) par sujets.

7.4 Annexe 4 – Effet du mapping sur la représentation mentale de la taille

Cette annexe présente une expérience purement perceptive qui s'est attachée à évaluer si l'on pouvait quand même évoquer des tailles différentes en modifiant le mapping, et si une relation entre la taille évoquée et le coefficient de mapping α pouvait être trouvée. Par exemple, en doublant le mapping α , évoque-t-on une forme deux fois plus grande ?

Méthode

Sujets. 18 sujets ont pris part à cette expérience (2 femmes) d'âge moyen 28 ans (SD=10.4). Aucun d'entre eux n'était familier avec le sujet de l'expérience.

Appareillage. Les sujets étaient assis devant un écran d'ordinateur où était présenté l'interface de test. Les sons étaient diffusés avec un casque Sennheiser HD-650. L'interface de test était programmée avec le logiciel Max/MSP.

Stimuli. 

- *Formes géométriques.* Une ellipse d'excentricité 0.9 de demi-grand axe 5 cm a été utilisée pour générer les profils de vitesse.
- *Profils de vitesse.* La dynamique du mouvement le long de l'ellipse a été définie selon le mouvement de Lissajous suivant (i.e. respectant la loi en puissance 1/3, voir 7.1) :

$$\begin{cases} x(t) = a \cos(\frac{2\pi}{T}t) \\ y(t) = b \sin(\frac{2\pi}{T}t) \end{cases} \quad (7.5)$$

où a et b sont respectivement le demi-grand et demi-petit axe. La période choisie fait 1.4 s.

- *Stimuli acoustiques.* Les sons de frottement ont été synthétisés avec le modèle phénoménologique présenté en annexe 7.2, et à partir des profils de vitesse $v_t(t)$ calculés à partir des mouvements qui viennent d'être définis. Dans le modèle de synthèse, le filtre passe-bande utilisé est un filtre Butterworth d'ordre 2 avec un facteur de qualité égal à 5. Le but de l'expérience est d'évaluer la relation entre le mapping et la taille de la forme évoquée. Dans la tâche qui va être décrite juste après, les sujets devaient donc manipuler le coefficient de mapping α entre la fréquence centrale du filtre passe-bande et la vitesse tangentielle du mouvement – $f_c(t) = \alpha v_t(t)$ pour évoquer des formes de tailles différentes.

Tâche et Procédure. Dans cette expérience les sujets devaient calibrer le mapping entre la vitesse tangentielle du mouvement et la fréquence centrale du filtre passe bande. Pour cela, un test d'ajustement a été mis en place, où les sujets écoutaient le son correspondant à un mapping de référence α_0 en ajustant l'autre son pour qu'il leur évoque une forme soit 2 fois plus petite, soit 2 fois plus grande, soit 4 fois plus grande que la forme évoquée par le son de référence. Trois mappings de référence ont été choisis : $\alpha_1 = 10, \alpha_2 = 20, \alpha_3 = 40$. Pour modifier le mapping, les sujets disposaient de 6 types de touches différentes : "< < <"; "< <"; "<" et ">"; "> >"; "> > >". En fonction de la touche, la valeur du mapping de l'ellipse était modifiée de façon plus ou moins importante, plus il y avait de flèches plus la modification était grande. Afin d'éviter un possible biais non-sensoriel, le rôle des flèches variait ainsi, les flèches *de droite* pouvaient diminuer la valeur de α dans un essai et l'augmenter

dans un autre. Une série de flèches (de gauche ou de droite selon l'essai) diminuait donc la valeur de α en la multipliant avec les valeurs d'un des trois jeux de paramètres suivants : $\{0.7, 0.85, 0.98\}$ ou $\{0.8, 0.9, 0.98\}$ ou $\{0.9, 0.95, 0.98\}$, et l'autre l'augmentait avec les facteurs multiplicatifs suivants : $\{1.02, 1.15, 1.3\}$ ou $\{1.02, 1.1, 1.2\}$ ou $\{1.02, 1.05, 1.1\}$, les plus petites valeurs correspondaient soit à la flèche "<" soit à la flèche ">". Le choix du jeu de valeurs utilisées pour les flèches et le rôle de chaque flèche (i.e. augmenter ou diminuer la valeur du mapping) était contrebalancé à travers les sujets. Pour chaque valeur de α_0 , les sujets devaient ajuster le mapping correspondant à trois cibles : 2 fois plus petite, 2 fois plus grande, ou 4 fois plus grande. Chaque couple (α_i) x (cible) a été répété 6 fois, soit au final 54 essais : 3 (mappings de référence) x 3 (cibles) x {6 répétitions}.

Analyse des données. Pour chaque mapping α_i , et chaque cible, 6 valeurs finales de α ont été collectées et moyennées. Afin de pouvoir comparer les résultats entre les différents mappings de références, la grandeur qui sera considérée pour les analyses sera le ratio entre la valeur finale et la valeur de référence $\rho_i = \frac{\alpha}{\alpha_i}$. Quel que soit le mapping de référence, les valeurs théoriques de ρ sont 0.5 pour la cible deux fois plus petite, 2 pour la cible deux fois plus grande, et 4 pour la cible 4 fois plus grande. Afin de comparer les différentes conditions, une ANOVA à mesures répétées à 2 facteurs (mapping de référence) x (cible) a été réalisée. Les effets significatifs ont été analysés avec un test post-hoc de Newman-Keuls.

Résultats

La table 7.7 et la figure 7.6 présentent les valeurs de ρ pour les différents mappings et cibles. Tout d'abord, l'ANOVA à mesures répétées a mis en évidence une influence de la cible sur la valeur de la cible calibrée ($F(2, 36) = 58.391, p < .001$). En particulier on peut voir que l'échelle du mapping est globalement bien cohérente avec l'échelle de la taille de la forme imaginée, par exemple, pour évoquer une forme 2 fois ou 4 fois plus grande, les sujets ont bien doublé le mapping ($\bar{\rho}_{2x} = 1.96$, CI95% [1.774 2.160] et $\bar{\rho}_{4x} = 3.81$, CI95% [3.005 4.630]). Seule la cible deux fois plus petite (i.e. 0.5x) a été légèrement surestimée en moyenne ($\bar{\rho}_{0.5x} = .56$, CI95% [.517 .618]). Cela montre que l'on peut bien évoquer des formes de tailles différentes en modifiant le mapping α et que de plus, la relation entre la fréquence du filtre et la taille moyenne est globalement linéaire. Par ailleurs, un effet du mapping est également mis en évidence ($F(2, 36) = 16.408, p < .001$). Cet effet révèle que pour le mapping $\alpha_1 = 10$, les valeurs de cibles ont été calibrées à des valeurs plus élevées que pour les deux autres mappings $\alpha_2 = 20$ et $\alpha_3 = 40$ ($p < .001$ pour les deux comparaisons).

Discussion

Dans cette expérience, les sujets devaient ajuster le mapping entre la vitesse tangentielle du mouvement et la fréquence centrale du filtre passe bande du modèle de synthèse de son de frottement afin d'évoquer des ellipses de tailles différentes. Les résultats ont mis en évidence que le mapping permettait bien d'évoquer des formes de tailles différentes. De plus, la relation entre le mapping et la taille évoquée est bien linéaire.

Ce résultat montre que le processus d'imagination d'une forme à partir d'un son est complètement différent du processus de synchronisation audio-moteur impliqué dans la l'expérience présentée en annexe 7.2. En effet, on a vu que le mapping n'influe pas de façon implicite sur la taille de l'ellipse dessinée alors que dans cette expérience purement perceptive, les sujets ont pu s'imaginer et calibrer des formes de tailles différentes en modifiant le mapping.

Ce paradigme est très intéressant notamment lorsque l'on envisage d'utiliser les sons pour l'apprentissage de geste expert. De nombreuses études ont montré l'intérêt de la simulation mentale

Cible	α	$\bar{\rho}$	CI95% _{inf}	CI95% _{sup}
0.5x	10	0.56	0.511	0.620
	20	0.59	0.535	0.657
	40	0.54	0.464	0.617
2x	10	2.29	1.983	2.614
	20	1.91	1.744	2.088
	40	1.68	1.501	1.873
4x	10	4.73	3.542	5.935
	20	3.58	2.888	4.282
	40	3.12	2.411	3.845

TABLE 7.7 – Moyennes et intervalles de confiance à 95% pour les valeurs finales de cibles calibrées en fonction des cibles théoriques et des différents mappings.

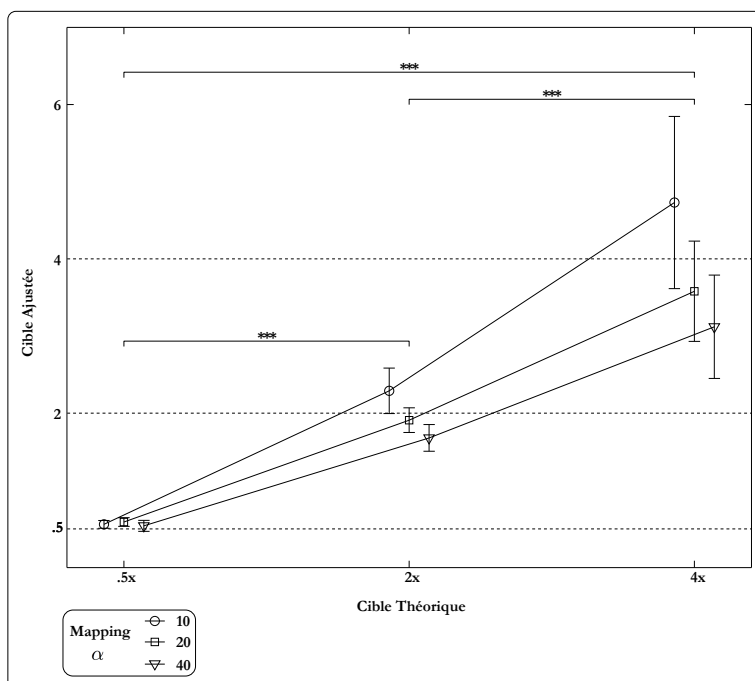



FIGURE 7.6 – Moyennes et intervalles de confiance à 95% pour les valeurs finales de cibles calibrées en fonction des cibles théoriques et des différents mappings. Les effets significatifs entre les différentes cibles correspondent à l'effet principal sur la cible (***: $p < .001$).

des actions pour l'apprentissage de gestes experts, comme chez les sportifs (Feltz et Landers, 1983; Pezzulo *et al.*, 2010; Bell *et al.*, 2009; Ramsey *et al.*, 2008), chez les danseurs (Calvo-Merino *et al.*, 2005), ou encore dans des processus de réhabilitation motrice Mulder (2007). Dans l'hypothèse où l'on souhaite faire apprendre la dynamique et la géométrie d'un mouvement à partir d'un son, cette expérience met en évidence que l'on peut évoquer des trajectoires de tailles différentes en jouant sur le rapport entre la vitesse et la fréquence. On peut en effet imaginer une tâche de simulation mentale de la dynamique d'un son correspondant à la dynamique d'un geste. On voit souvent les sportifs de haut niveau comme les skieurs répéter leurs actions avant de s'élancer, et on pourrait les aider en les faisant simuler la dynamique du son plutôt que la dynamique visuelle.

7.5 Annexe 5 – Autres exemples de friction non linéaire

Cette annexe présente d'autres exemples de friction non linéaire que celui du verre qui couine et qui chante présenté dans le chapitre 5.

- **Porte qui grince.**  La figure 7.7 présente la transformée de Gabor de l'enregistrement d'une porte qui grince. L'analyse de l'harmonicit   montre que cet extrait sonore est bien harmonique, et que sa fr  quence fondamentale f_0 varie au cours du temps. L'  volution de f_0 n'est pas lin  aire et on peut faire l'hypoth  se que cela est d      au fait que la vitesse de rotation et la pression au niveau de l'axe de la porte ne sont pas constants au cours du temps. Les transitions soudaines entre les diff  rents   tats vibratoires sont clairement visibles sur le diagramme temps-fr  quence. Une caract  ristique de la morphologie sonore associ  e au grincement de porte est le lien entre les fortes variations de la fr  quence fondamentale f_0 et la quantit   du bruit : lorsque f_0 est faible, le signal est tr  s bruit   et quand f_0 est   lev  e le son est tr  s harmonique, ce qui entraine de grandes variations de timbre.

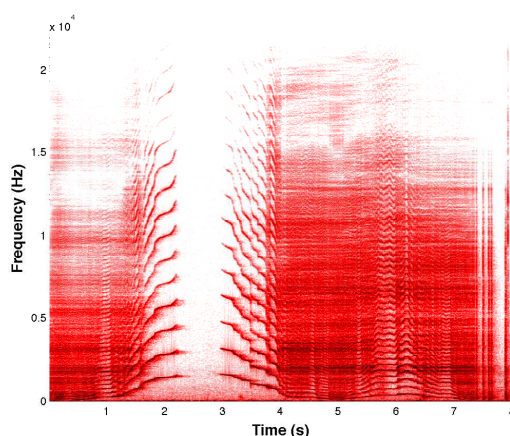



FIGURE 7.7 – Transform  e de Gabor du son produit par une porte qui grince – Figure issue de Thoret et al. (b)

- **Couinement sur une assiette humide.**  La figure 7.8 pr  sente la transform  e de Gabor de l'enregistrement d'une personne qui fait couiner une assiette mouill  e. On constate que la variation de la fr  quence fondamentale est beaucoup plus faible que dans le cas de la porte qui grince. De plus, on peut constater que cette fr  quence fondamentale varie lentement autour d'une valeur centrale au cours du temps. Les modes de l'assiette sont   galement clairement visibles et sont excit  s quand la valeur de la fr  quence fondamentale f_0 est proche. Cependant, f_0 ne se bloque jamais sur l'un deux. Cette observation va dans le sens du mod  le source-r  sonance et du paradigme action-objet.

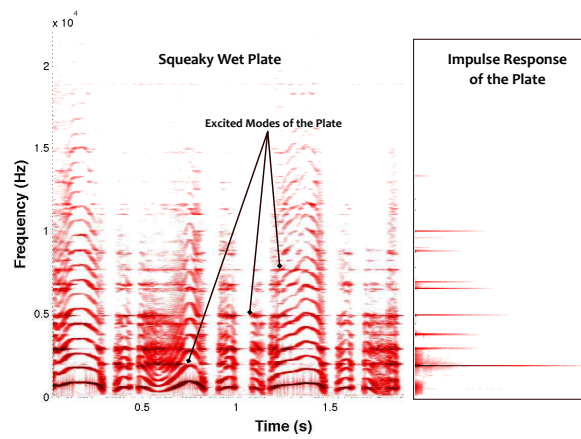


FIGURE 7.8 – Transformée de Gabor du son produit par un doigt qui fait couiner une assiette humide
– Figure issue de Thoret *et al.* (b)

7.6 Annexe 6 – Publications et communications associées à la thèse

Les publications et communications associées à la thèse sont séparées en deux parties : la première avec les publications en premier auteur concernant les travaux directement réalisés et la seconde avec les publications réalisées dans le cadre des collaborations réalisées au cours de la thèse :

Publications et communications du premier auteur :

Journaux à comité de lecture (1)

1. **Thoret, E.**, Aramaki, M., Kronland-Martinet, R., Velay, J. L., & Ystad, S. (2014). From Sound to Shape : Auditory Perception of Drawing Movements, *Journal of Experimental Psychology : Human Perception and Performance*, 40, 983-994. doi : 10.1037/a0035441

Chapitres d'ouvrage (2)

2. **Thoret, E.**, Aramaki, M., Bourdin C., Bringoux L., Kronland-Martinet R., & Ystad S. (2014). Audio-motor synchronization : the effect of mapping between kinematics and acoustic cues on geometric motor features. In *Sound, Music and Motion*, LNCS 8905, pp234-245, Springer Berlin Heidelberg. doi : 10.1007/978-3-319-12976-1_15

1. **Thoret, E.**, Aramaki, M., Kronland-Martinet, R., Velay, J. L., & Ystad, S. (2013). Reenacting sensorimotor features of drawing movements from friction sounds. In *From Sounds to Music and Emotions*, LNCS vol 7900, 130-153. Springer Berlin Heidelberg. doi : 10.1007/978-3-642-41248-6_8

Communications dans des conférences internationales à comité de lecture (7)

7. **Thoret, E.**, Aramaki, M., Bringoux L., Kronland-Martinet R., & Ystad S. (2014). When acoustic stimuli turn visual circles into ellipses : sounds evoking accelerations modify visuo-motor coupling, *15th International Multisensory Research Forum (IMRF)*, Amsterdam, Jun. 11-14, 2014 – Poster

6. **Thoret, E.**, Aramaki, M., Kronland-Martinet, R., Velay, J. L., & Ystad, S. (2014). Auditory Perception of Biological Movements : an evidence of cognitive specificities from sound synthesis, *15th International Multisensory Research Forum (IMRF)*, Amsterdam, Jun. 11-14, 2014 – Poster

5. **Thoret, E.**, Aramaki, M., Bourdin C., Bringoux L., Kronland-Martinet R., & Ystad S. (2013). Synchronizing Gestures with Friction Sounds : Work in Progress, *Sound, Music and Motion, Proceedings of the 10th International Symposium on Computer Music Multidisciplinary Research (CMMR)*, 15-18 Oct. 2013, Marseille – Oral presentation

4. **Thoret E.**, Aramaki M., Gondre C., Kronland-Martinet R., & Ystad S. (2013). Controlling a Non Linear Friction Model for Evocative Sound Synthesis Applications, *Proceedings of the 16th International Conference on Digital Audio Effects (DAFx)*, Maynooth, Ireland, Sept. 2013 – Oral presentation

3. **Thoret E.**, Aramaki M., Kronland-Martinet R., Velay J.L., & Ystad S. (2012). From Shape to Sound : sonification of two dimensional curves by reenaction of biological movements, *Music and Emotions, Proceedings of the 9th International Symposium on Computer Music Modeling and Retrieval (CMMR)*, 19-22 June, London, 2012 – Oral presentation

2. **Thoret E.**, Aramaki M., Kronland-Martinet R., Velay J.L., & Ystad S. (2012). Sonifying drawings : characterization of perceptual attributes of sounds produced by human gestures, *Proceedings of Acoustics 2012*, 23-27 April, Nantes, France, April 2012 – Oral presentation

1. **Thoret E.**, Aramaki M., Kronland-Martinet R., Velay J.L., & Ystad S. (2011). Sonification of drawings by virtually reenacting biological movements, *Versatile Sound Models for Interaction in Audio-Graphic Virtual Environments : Control of Audio-graphic Sound Synthesis Satellite Workshop at the 14th International Conference on Digital Audio Effects (DAFx)*, Paris, 2011 – Oral presentation

Séminaires invités (1)

1. **Thoret, E.** (2014) Auditory perception of drawing movements : when motor cues shape auditory cognitive processes, *Legos days*, Ircam, March 18, 2014, Paris

Conférences nationales (1)

1. **Thoret, E.** (2012) Approche éactive des relations son/geste pour le contrôle de la synthèse sonore et le guidage. *2èmes Journées Perception Sonore - SFA*, 10–11 Dec. 2012, Marseille – Poster

Collaborations :

Journaux à comité de lecture (2)

2. Danna, J., Fontaine, M., Paz-Villagran, V., Gondre, C., **Thoret, E.**, Aramaki, M., Kronland-Martinet, R., Ystad, S., & Velay, J.-L. (in press). The effect of real-time auditory feedback on learning new characters. *Human Movement Science*. doi : 10.1016/j.humov.2014.12.002

1. Conan, S., **Thoret, E.**, Aramaki, M., Derrien, O., Gondre, C., Kronland-Martinet, R., & Ystad, S. (2014). An Intuitive Synthesizer of Continuous Interaction Sounds : Rubbing, Scratching and Rolling. *Computer Music Journal*, 38(4), 24-37. doi : 10.1162/COMJ_a_00266

Chapitres d'ouvrage (1)

1. Danna J., Velay J.-L., Paz-Villagran V., Capel A., Pétrou C., Gondre C., **Thoret E.**, Aramaki M., Ystad S., & Kronland-Martinet, R. (2014) Movement sonification for the diagnosis and the rehabilitation of graphomotor disorders. In *Sound, Music and Motion*, LNCS 8905, pp246-255, Springer Berlin Heidelberg. doi : 10.1007/978-3-319-12976-1_16

Communications dans des conférences internationales à comité de lecture (4)

4. Danna J., Velay J.-L., Paz-Villagran V., Capel A., Pétrou C., Gondre C., **Thoret E.**, Aramaki M., Ystad S., & Kronland-Martinet, R. (2013). Handwriting Movement Sonification for the Rehabilitation of Dysgraphia, *Sound, Music and Motion, Proceedings of the 10th International Symposium on Computer Music Multidisciplinary Research (CMMR)*, 15-18 Oct. 2013, Marseille – Oral presentation

3. Conan S., **Thoret E.**, Aramaki M., Derrien O., Gondre C., Kronland-Martinet R., & Ystad S. (2013). Navigating in a Space of Synthesized Interactions-Sounds : Rubbing, Scratching and Rolling, *Proceedings of the 16th Conference on Digital Audio Effects (DAFx)*, Maynooth, Ireland, Sept. 2013 – Oral presentation, *Best paper award in sound synthesis*

2. Danna J., Paz-Villagran V., **Thoret E.**, Gondre C., Kronland-Martinet R., Capel A., Pétrou C., Pierre P., Cazes O., Limozin A., & Velay J.-L. (2013). Sonifying handwriting movements as real-time auditory feedback for the rehabilitation of dysgraphia. *Abstract in Progress in Motor Control IX*, Montreal, July 13-16 – Poster

1. Conan S., Aramaki M., Kronland-Martinet R., **Thoret E.** & Ystad S (2012). Perceptual differences between sounds produced by different continuous interactions, *Proceedings of Acoustics 2012*, 23-27 April Nantes, France, April 2012 – Poster

Conférences nationales (2)

2. Danna, J., Fontaine, M., Gondre, C., **Thoret, E.**, Aramaki, M., Kronland-Martinet, R., Ystad, S., & Velay, J.-L. (2013). Sonifier le geste d'écriture peut-il faciliter son apprentissage ?, 15ème Congrès International de l'ACAPS, Grenoble, France, 29-31 Octobre 2013 – Poster

1. Danna, J., Paz-Villagran, V., Velay, J.-L., Gondre, C., Kronland-Martinet, R., Ystad, S., Aramaki, M., **Thoret, E.**, Conan, S., Voinier, T., Omer, H., & Torrèsani, B. (2012). Sonifier l'écriture : un outil pour le diagnostic et la remédiation de la dysgraphie, IIIèmes Journée Scientifique du Centre de Référence des Troubles d'Apprentissage (CERTA)/RESODYS, Marseille, May 25, 2012 – Oral Presentation

Autres (2)

2. Conan, S., **Thoret, E.**, Gondre, C., Aramaki, M., Kronland-Martinet, R., & Ystad, S. (2013). An Intuitive Synthesizer of Sustained Interaction Sounds. *Sound, Music and Motion, Proceedings of the 10th International Symposium on Computer Music Multidisciplinary Research (CMMR)*, 15-18 Oct., Marseille – Demonstration

1. Danna, J., Paz-Villagran, V., Velay, J.-L., Gondre, C., Kronland-Martinet, R., Ystad, S., Aramaki, M., **Thoret, E.**, Conan, S., Voinier, T., Omer, H., & Torrèsani, B. (2012). Sonifier l'écriture : un outil pour le diagnostic et la remédiation de la dysgraphie, *Développements*, 12, 2012, 32-40

Bibliographie

1. AKAY, A. (2002). Acoustics of friction. *The Journal of the Acoustical Society of America*, 111(4):1525–1548.
2. ARAMAKI, M., BESSON, M., KRONLAND-MARTINET, R. et YSTAD, S. (2011). Controlling the perceived material in an impact sound synthesizer. *IEEE Transactions on Audio, Speech, and Language Processing*, 19(2):301–314.
3. ARAMAKI, M., GONDRE, C., KRONLAND-MARTINET, R., VOINIER, T. et YSTAD, S. (2010). Imagine the sounds : An intuitive control of an impact sound synthesizer. In YSTAD, S., ARAMAKI, M., KRONLAND-MARTINET, R. et JENSEN, K., éditeurs : *Auditory Display*, volume 5954 de *Lecture Notes in Computer Science*, pages 408–421. Springer.
4. ARAMAKI, M. et KRONLAND-MARTINET, R. (2006). Analysis-synthesis of impact sounds by real-time dynamic filtering. *IEEE Transactions on Audio, Speech, and Language Processing*, 14(2):695 – 705.
5. ARNOTT, S. R., BINNS, M. A., GRADY, C. L. et ALAIN, C. (2004). Assessing the auditory dual-pathway model in humans. *Neuroimage*, 22(1):401–408.
6. ATHENES, S., SALLAGOÏTY, I., ZANONE, P.-G. et ALBARET, J.-M. (2004). Evaluating the coordination dynamics of handwriting. *Human movement science*, 23(5):621–641.
7. AVANZINI, F., SERAFIN, S. et ROCCHESO, D. (2005). Interactive simulation of rigid body interaction with friction-induced sound generation. *IEEE Transactions on Speech and Audio Processing*, 13(5):1073–1081.
8. BANGERT, M., PESCHEL, T., SCHLAUG, G., ROTTE, M., DRESCHER, D., HINRICHS, H., HEINZE, H.-J. et ALTENMÜLLER, E. (2006). Shared networks for auditory and motor processing in professional pianists : evidence from fmri conjunction. *Neuroimage*, 30(3):917–926.
9. BARSALOU, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and brain sciences*, 22(04):637–660.
10. BELL, R. J., SKINNER, C. H. et FISHER, L. A. (2009). Decreasing putting yips in accomplished golfers via solution-focused guided imagery : A single-subject research design. *Journal of Applied Sport Psychology*, 21(1):1–14.
11. BEN ABDELOUNIS, H., LE BOT, A., PERRET-LIAUDET, J. et ZAHOUANI, H. (2010). An experimental study on roughness noise of dry rough flat surfaces. *Wear*, 268(1–2):335–345.
12. BERGER, K. W. (1964). Some factors in the recognition of timbre. *The Journal of the Acoustical Society of America*, 36(10): 1888–1891.
13. BERTHOZ, A. (1997). *Le sens du mouvement*. Odile Jacob.
14. BESSON, P., BOURDIN, C. et BRINGOUX, L. (2011). A comprehensive model of audiovisual perception : Both percept and temporal dynamics. *PloS one*, 6(8):e23811.
15. BIDET-ILDEI, C., ORLIAGUET, J.-P., SOKOLOV, A. N. et PAVLOVA, M. (2006). Perception of elliptic biological motion. *Perception*, 35:1137–1147.
16. BIESS, A. (2013). Shaping of arm configuration space by prescription of non-euclidean metrics with applications to human motor control. *Physical Review E*, 87(1):012729–1,012729–15.
17. BIGAND, E., VIEILLARD, S., MADURELL, F., MAROZEAU, J. et DACQUET, A. (2005). Multidimensional scaling of emotional responses to music : The effect of musical expertise and of the duration of the excerpts. *Cognition & Emotion*, 19(8):1113–1139.

18. BINET, A. et COURTIER, J. (1893). Sur la vitesse des mouvements graphiques. *Revue Philosophique de la France et de l'Étranger*, 35:664–671.
19. CALVO-MERINO, B., GLASER, D., GRÈZES, J., PASSINGHAM, R. et HAGGARD, P. (2005). Action observation and acquired motor skills : an fmri study with expert dancers. *Cerebral cortex*, 15(8):1243–1249.
20. Canudas de WIT, C., OLSSON, H., ASTRÖM, K. J. et LISCHINSKY, P. (1995). A new model for control of systems with friction. *IEEE Transactions on automatic control*, 40(3):419–425.
21. CARAMIAUX, B. (2011). *Etudes sur la relation geste-son en performance musicale*. Thèse de doctorat, Université Pierre et Marie Curie - Paris VI.
22. CARELLO, C., ANDERSON, K. L. et KUNKLER-PECK, A. J. (1998). Perception of object length by sound. *Psychological Science*, 9(3):211–214.
23. CARLILE, S. et BEST, V. (2002). Discrimination of sound source velocity in human listeners. *The Journal of the Acoustical Society of America*, 111(2):1026–1035.
24. CARLYON, R. P., MACHEREY, O., FRIJNS, J. H. M., AXON, P. R., KALKMAN, R. K., BOYLE, P., BAGULEY, D. M., BRIGGS, J., DEEKS, J. M. et BRIAIRE, J. J. (2010). Pitch comparisons between electrical stimulation of a cochlear implant and acoustic stimuli presented to a normal-hearing contralateral ear. *Journal of the Association for Research in Otolaryngology*, 11(4):625–640.
25. CASILE, A. et GIESE, M. A. (2006). Nonvisual motor training influences biological motion perception. *Current Biology*, 16(1):69–74.
26. CHANGEUX, J.-P. (1983). *L'homme neuronal*. Fayard Paris.
27. CHAO, L. L. et MARTIN, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *Neuroimage*, 12(4):478–484.
28. CHOWNING, J. M. (1971). The simulation of moving sound sources. *Journal of the Audio Engineering Society*, 19(1):2–6.
29. CISEK, P. et KALASKA, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience*, 33(1):269–298.
30. CONAN, S. (2014). *Contrôle intuitif de la synthèse sonore d'interactions solidiennes : vers les métaphores sonores*. Thèse de doctorat, Ecole Centrale Marseille.
31. CONAN, S., ARAMAKI, M., KRONLAND-MARTINET, R., THORET, E. et YSTAD, S. (2012). Perceptual differences between sounds produced by different continuous interactions. *Proceedings of Acoustics 2012, Nantes*, pages 581–586.
32. CONAN, S., DERRIEN, O., ARAMAKI, M., YSTAD, S. et KRONLAND-MARTINET, R. (2014a). A synthesis model with intuitive control capabilities for rolling sounds. *IEEE/ACM Transactions on Speech, Audio and Language Processing*, 22(8):1260–1273.
33. CONAN, S., THORET, E., ARAMAKI, M., DERRIEN, O., GONDRE, C., KRONLAND-MARTINET, R. et YSTAD, S. (2013). Navigating in a space of synthesized interaction-sounds : Rubbing, scratching and rolling sounds. In *Proceedings of the 16th International Conference on Digital Audio Effects (DAFx), Maynooth, Ireland, 2-5 Sept. 2013*, pages 202–209.
34. CONAN, S., THORET, E., ARAMAKI, M., DERRIEN, O., GONDRE, C., YSTAD, S. et KRONLAND-MARTINET, R. (2014b). An intuitive synthesizer of continuous interaction sounds : Rubbing, scratching and rolling. *Computer Music Journal*, 38(4):24–37.
35. COOK, R., GAULE, A., AICHELBURG, C. et PRESS, C. (2014). A lexical analysis of environmental sound categories. *Journal of Experimental Psychology : Human Perception and Performance*. Advance publication online.
36. COULOMB, C. A. (1821). *Théorie des machines simples*.
37. CRAIG, C. M., DELAY, D., GREALLY, M. A. et LEE, D. N. (2000). Guiding the swing in golf putting. *Nature*, 405(6784):295–296.

38. CREEM-REGEHR, S. H. et LEE, J. N. (2005). Neural representations of graspable objects : are tools special ? *Cognitive Brain Research*, 22(3):457–469.
39. DANNA, J. (2011). *Dynamique de coordination dans la formation de la trace écrite chez l'adulte et l'enfant*. Thèse de doctorat, Université de Toulouse, Université Toulouse III-Paul Sabatier.
40. DANNA, J., ATHÈNES, S. et ZANONE, P.-G. (2011). Coordination dynamics of elliptic shape drawing : Effects of orientation and eccentricity. *Human movement science*, 30(4):698–710.
41. DANNA, J., FONTAINE, M., PAZ-VILLAGRÁN, V., GONDRE, C., THORET, E., ARAMAKI, M., KRONLAND-MARTINET, R., YSTAD, S. et VELAY, J.-L. (In press). The effect of real-time auditory feedback on learning new characters. *Human Movement Science*.
42. DANNA, J., PAZ-VILLAGRÁN, V. et VELAY, J.-L. (2013a). Signal-to-noise velocity peaks difference : A new method for evaluating the handwriting movement fluency in children with dysgraphia. *Research in developmental disabilities*, 34(12):4375–4384.
43. DANNA, J., VELAY, J., PAZ-VILLAGRÁN, V., A., C., PÉTROZ, C., GONDRE, C., THORET, E., ARAMAKI, M., YSTAD, S. et KRONLAND-MARTINET, R. (2013b). Handwriting movement sonification for the rehabilitation of dysgraphia. In *Proceedings of the 10th International Symposium on Computer Music Multidisciplinary Research (CMMR)*, 15-18 Oct., Marseille, pages 200–208.
44. DAYAN, E., CASILE, A., LEVIT-BINNUN, N., GIESE, M. A., HENDLER, T. et FLASH, T. (2007). Neural representations of kinematic laws of motion : evidence for action-perception coupling. *Proceedings of the National Academy of Sciences*, 104(51):20582–20587.
45. DE SAUSSURE, F. (1916). *Cours de linguistique générale*, volume 1. Otto Harrassowitz Verlag.
46. de SPERATI, C. et VIVIANI, P. (1997). The relationship between curvature and velocity in two-dimensional smooth pursuit eye movements. *The Journal of Neuroscience*, 17(10):3932–3945.
47. DECETY, J., JEANNEROD, M. et PRABLANC, C. (1989). The timing of mentally represented actions. *Behavioural brain research*, 34(1):35–42.
48. DECETY, J. et MICHEL, F. (1989). Comparative analysis of actual and mental movement times in two graphic tasks. *Brain and cognition*, 11(1):87–97.
49. DESCARTES, R. (1972). *Traité de l'homme. Oeuvres de Descartes*, 11:119–215.
50. DE'SPERATI, C. et STUCCHI, N. (1995). Visual tuning to kinematics of biological motion : the role of eye movements. *Experimental Brain Research*, 105(2):254–260.
51. DEVEREUX, G. et SINACEUR, H. (1967). *De l'angoisse à la méthode dans les sciences du comportement*. Aubier.
52. DOUNSKAIA, N., VAN GEMMERT, A. W. A. et STELMACH, G. E. (2000). Interjoint coordination during handwriting-like movements. *Experimental Brain Research*, 135(1):127–140.
53. ELLIS, D. P. W. (2003). Sinewave and sinusoid+noise analysis/synthesis in matlab. <http://www.ee.columbia.edu/~dpwe/resources/matlab/sinemodel/>.
54. FARNELL, A. (2010). *Designing sound*. MIT Press Cambridge.
55. FELDMAN, J. (2000). Bias toward regular form in mental shape spaces. *Journal of Experimental Psychology : Human Perception and Performance*, 26(1):152.
56. FELTZ, D. L. et LANDERS, D. M. (1983). The effects of mental practice on motor skill learning and performance : A meta-analysis. *Journal of sport psychology*, 5(1):25–57.
57. FLASH, T. et HANDZEL, A. A. (2007). Affine differential geometry analysis of human arm movements. *Biological cybernetics*, 96(6):577–601.
58. FRAISSE, P., OLÉRON, G. et PAILLARD, J. (1953). Les effets dynamogéniques de la musique. étude expérimentale. *L'année psychologique*, 53(1):1–34.

59. FRAISSE, P., PICHOT, P. et CLAIROUIN, G. (1949). Les aptitudes rythmiques. étude comparée des oligophrènes et des enfants normaux. *Journal de Psychologie Normale et Pathologique*, 42:309–330.
60. FREITAG, C. M., KONRAD, C., HÄBERLEN, M., KLESER, C., von GONTARD, A., REITH, W., TROJE, N. F. et KRICK, C. (2008). Perception of biological motion in autism spectrum disorders. *Neuropsychologia*, 46(5):1480–1494.
61. FREYD, J. J. (1983a). The mental representation of movement when static stimuli are viewed. *Perception & Psychophysics*, 33(6):575–581.
62. FREYD, J. J. (1983b). Representing the dynamics of a static form. *Memory & Cognition*, 11(4):342–346.
63. FRÉMIOT, M., MANDELBROJT, J., FORMOSA, M., DELALANDE, G., PEDLER, E., MALBOSC, P. et GOBIN, P. (1996). Les unités sémiotiques temporelles—Éléments nouveaux d'analyse musicale. *Diffusion ESKA. MIM Laboratoire Musique et Informatique de Marseille (MIM), documents musurgia édition*, 13.
64. GALLESE, V., FADIGA, L., FOGASSI, L. et RIZZOLATTI, G. (1996). Action recognition in the premotor cortex. *Brain*, 119(2):593–609.
65. GANDEMER, L., PARSEIHIAN, G., KRONLAND-MARTINET, R. et BOURDIN, C. (2014). The influence of horizontally rotating sound on standing balance. *Experimental brain research*, 232(12):3813–3820.
66. GAVER, W. (1993a). How do we hear in the world ? explorations in ecological acoustics. *Ecological psychology*, 5(4):285–313.
67. GAVER, W. W. (1993b). What in the world do we hear ? an ecological approach to auditory event perception. *Ecological Psychology*, 5(1):1–29.
68. GIBSON, J. J. (1966). *The senses considered as perceptual systems*. Houghton Mifflin.
69. GIBSON, J. J. (1979). *The ecological approach to visual perception*. Psychology Press.
70. GIESE, M. A. et POGGIO, T. (2003). Neural mechanisms for the recognition of biological movements. *Nature Reviews Neuroscience*, 4(3):179–192.
71. GIORDANO, B. L. et MCADAMS, S. (2006). Material identification of real impact sounds : Effects of size variation in steel, glass, wood, and plexiglass plates. *The Journal of the Acoustical Society of America*, 119(2):1171–1181.
72. GIORDANO, B. L., MCADAMS, S., ZATORRE, R. J., KRIEGESKORTE, N. et BELIN, P. (2012). Abstract encoding of auditory objects in cortical activity patterns. *Cerebral Cortex*, page bhs162.
73. GIORDANO, B. L., ROCCHESO, D. et MCADAMS, S. (2010). Integration of acoustical information in the perception of impacted sound sources : The role of information accuracy and exploitability. *Journal of Experimental Psychology : Human Perception and Performance*, 36(2):462–476.
74. GOLDSTEIN, H. (1962). *Classical mechanics*, volume 4. Pearson Education India.
75. GRAFTON, S. T., FADIGA, L., ARBIB, M. A. et RIZZOLATTI, G. (1997). Premotor cortex activation during observation and naming of familiar tools. *Neuroimage*, 6(4):231–236.
76. GRANTHAM, D. W. (1986). Detection and discrimination of simulated motion of auditory targets in the horizontal plane. *The Journal of the Acoustical Society of America*, 79(6):1939–1949.
77. GRASSI, M. (2005). Do we hear size or sound ? balls dropped on plates. *Perception & Psychophysics*, 67(2):274–284.
78. GRASSI, M., PASTORE, M. et LEMAITRE, G. (2013). Looking at the world with your ears : How do we get the size of an object from its sound ? *Acta Psychologica*, 143(1):96–104.
79. GREENWALD, A. G. (1970). A choice reaction time test of ideomotor theory. *Journal of Experimental Psychology*, 86(1):20–25.
80. GREY, J. M. et GORDON, J. W. (1978). Perceptual effects of spectral modifications on musical timbres. *The Journal of the Acoustical Society of America*, 63(5):1493–1500.

81. GRÈZES, J. et DECETY, J. (2002). Does visual perception of object afford action ? evidence from a neuroimaging study. *Neuropsychologia*, 40(2):212–222.
82. GREZES, J., FONLUPT, P., BERTENTHAL, B., DELON-MARTIN, C., SEGEBARTH, C. et DECETY, J. (2001). Does perception of biological motion rely on specific brain regions ? *Neuroimage*, 13(5):775–785.
83. GRIFFITHS, T. D. et WARREN, J. D. (2004). What is an auditory object ? *Nature Reviews Neuroscience*, 5(11):887–892.
84. GROSSMAN, E., DONNELLY, M., PRICE, R., PICKENS, D., MORGAN, V., NEIGHBOR, G. et BLAKE, R. (2000). Brain areas involved in perception of biological motion. *Journal of cognitive neuroscience*, 12(5):711–720.
85. GYGI, B. et SHAFIRO, V. (2006). General functions and specific applications of environmental sound research. *Frontiers in bioscience : a journal and virtual library*, 12:3152–3166.
86. HAKEN, H. (1983). *Advanced synergetics*. Springer.
87. HAKEN, H., KELSO, J. A. S. et BUNZ, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological cybernetics*, 51(5):347–356.
88. HAMSTRA-BLETZ, L. et BLÖTE, A. W. (1993). A longitudinal study on dysgraphic handwriting in primary school. *Journal of Learning Disabilities*, 26(10):689–699.
89. HARRIS, J. D. et SERGEANT, R. L. (1971). Monaural/binaural minimum audible angles for a moving sound source. *Journal of Speech, Language, and Hearing Research*, 14(3):618–629.
90. HAUEISEN, J. et KNÖSCHE, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of cognitive neuroscience*, 13(6):786–792.
91. HIDAKA, S., TERAMOTO, W., SUGITA, Y., MANAKA, Y., SAKAMOTO, S. et SUZUKI, Y. (2011). Auditory motion information drives visual motion perception. *PLoS One*, 6(3):e17499.
92. HOLLERBACH, J. M. (1981). An oscillation theory of handwriting. *Biological Cybernetics*, 39(2):139–156.
93. HOMMEL, B. (2004). Event files : feature binding in and across perception and action. *Trends in Cognitive Sciences*, 8(11):494–500.
94. HOMMEL, B., MÜSSELER, J., ASCHERSLEBEN, G. et PRINZ, W. (2001). The theory of event coding (TEC) : a framework for perception and action planning. *Behavioral and Brain Sciences*, 24(5):849–878.
95. HOUBEN, M. M. J., KOHLRAUSCH, A. et HERMES, D. J. (2004). Perception of the size and speed of rolling balls by sound. *Speech communication*, 43(4):331–345.
96. HOUBEN, M. M. J., KOHLRAUSCH, A. et HERMES, D. J. (2005). The contribution of spectral and temporal information to the auditory perception of the size and speed of rolling balls. *Acta Acustica united with Acustica*, 91(6):1007–1015.
97. HOUIX, O., LEMAITRE, G., MISDARIIS, N., SUSINI, P. et URDAPILLETA, I. (2012). A lexical analysis of environmental sound categories. *Journal of Experimental Psychology : Applied*, 18(1):52–80.
98. ISO-25178 (2010). Norme iso-25178 – spécification géométrique des produits (gps) – État de surface : Surfaique – partie 2 : Termes, définitions et paramètres d'états de surface.
99. JACK, C. E. et THURLOW, W. R. (1973). Effects of degree of visual association and angle of displacement on the "ventriloquism" effect. *Perceptual and motor skills*, 37(3):967–979.
100. JACK, W. R. (1895). The analysis of voluntary muscular movements by certain new instruments. *Journal of anatomy and physiology*, 29(4):473.
101. JAMES, K. H. et GAUTHIER, I. (2006). Letter processing automatically recruits a sensory–motor brain network. *Neuropsychologia*, 44(14):2937–2949.
102. JEANNEROD, M. (1995). Mental imagery in the motor context. *Neuropsychologia*, 33(11):1419–1432.
103. JEANNEROD, M. (2011). *La fabrique des idées*. Odile Jacob.

104. JOHANSSON, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & psychophysics*, 14(2):201–211.
105. KALÉNINE, S., CHEAM, C., IZARD, V. et GENTAZ, E. (2013). Adults and 5-year-old children draw rectangles and triangles around a prototype but not in the golden ratio. *British Journal of Psychology*, 104(3):400–412.
106. KALENINE, S., PINET, L. et GENTAZ, E. (2011). The visual and visuo-haptic exploration of geometrical shapes increases their recognition in preschoolers. *International Journal of Behavioral Development*, 35(1):18–26.
107. KELSO, J. A. S. (1984). Phase transitions and critical behavior in human bimanual coordination. *American Journal of Physiology*, 246(6 Pt 2):R1000–R1004.
108. KELSO, J. A. S. (1986). Pattern formation in speech and limb movements involving many degrees of freedom. *Experimental Brain Research*, 15:105–128.
109. KELSO, J. A. S. (1997). *Dynamic patterns : The self-organization of brain and behavior*. MIT press.
110. KELSO, J. A. S. (2012). Multistability and metastability : understanding dynamic coordination in the brain. *Philosophical Transactions of the Royal Society B : Biological Sciences*, 367(1591):906–918.
111. KELSO, J. A. S., HOLT, K. G., RUBIN, P. et KUGLER, P. N. (1981). Patterns of human interlimb coordination emerge from the properties of non-linear, limit cycle oscillatory processes : theory and data. *Journal of motor behavior*, 13(4):226–261.
112. KELSO, J. A. S., VATIKIOTIS-BATESON, E., SALTZMAN, E. L. et KAY, B. (1985). A qualitative dynamic analysis of reiterant speech production : Phase portraits, kinematics, and dynamic modeling. *The Journal of the Acoustical Society of America*, 77(1):266–280.
113. KENNEL, C., PIZZERA, A., HOHMANN, T., SCHUBOTZ, R. I., MURGIA, M., AGOSTINI, T. et RAAB, M. (2014). The perception of natural and modulated movement sounds. *Perception*, 43:796–804.
114. KEOGH, E. et RATANAMAHATANA, C. A. (2005). Exact indexing of dynamic time warping. *Knowledge and information systems*, 7(3):358–386.
115. KEPLER, J. (1604). *Paralipomènes à Vitellion*. Librairie Philosophique J. Vrin.
116. KITAGAWA, N. et ICHIHARA, S. (2002). Hearing visual motion in depth. *Nature*, 416(6877):172–174.
117. KLATZKY, R. L., PAI, D. K. et KROTKOV, E. P. (2000). Perception of material from contact sounds. *Presence : Teleoperators and Virtual Environments*, 9(4):399–410.
118. KLIN, A., LIN, D. J., GORRINDO, P., RAMSAY, G. et JONES, W. (2009). Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature*, 459(7244):257–261.
119. KOHLER, E., KEYSERS, C., UMITLA, M. A., FOGASSI, L., GALLESE, V. et RIZZOLATTI, G. (2002). Hearing sounds, understanding actions : action representation in mirror neurons. *Science*, 297(5582):846–848.
120. KÖHLER, W. (1929). *Gestalt psychology : An introduction to new concepts in modern psychology*. W.W. Norton & Company.
121. KONZEM, P. A. B. (1987). *Extended practice and patterns of bimanual interference*. Thèse de doctorat, Unpublished doctoral dissertation. University of Southern California.
122. KRONLAND-MARTINET, R. et VOINIER, T. (2008). Real-time perceptual simulation of moving sources : Application to the leslie cabinet and 3d sound immersion. *EURASIP J. Audio Speech Music Process.*, 2008:7 :1–7 :10.
123. KUNKLER-PECK, A. J. (2000). Hearing shape. *Journal of Experimental Psychology : Human Perception and Performance*, 26(1):279–294.
124. LA SCALEIA, B., ZAGO, M., MOSCATELLI, A., LACQUANITI, F. et VIVIANI, P. (2014). Implied dynamics biases the visual perception of velocity. *PloS one*, 9(3):e93020.
125. LACAN, J. (1966). Le séminaire sur la lettre volée. *Écrits*, 1:19–75.
126. LACQUANITI, F., TERZUOLO, C. et VIVIANI, P. (1983). The law relating the kinematic and figural aspects of drawing movements. *Acta Psychologica*, 54(1-3):115–130.

127. LAGRANGE, M., SCAVONE, G. et DEPALLE, P. (2010). Analysis/synthesis of sounds generated by sustained contact between rigid objects. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(3):509–518.
128. LAKATOS, S., MCADAMS, S. et CAUSSÉ, R. (1997). The representation of auditory source characteristics : Simple geometric form. *Perception & psychophysics*, 59(8):1180–1190.
129. LEE, D. N. (1998). Guiding movement by coupling taus. *Ecological psychology*, 10(3-4):221–250.
130. LEMAITRE, G. et HELLER, L. M. (2012). Auditory perception of material is fragile while action is strikingly robust. *The Journal of the Acoustical Society of America*, 131(2):1337–1348.
131. LEMAN, M. (2008). *Embodied music cognition and mediation technology*. Mit Press.
132. LEVIT-BINNUN, N., SCHECHTMAN, E. et FLASH, T. (2006). On the similarities between the perception and production of elliptical trajectories. *Experimental brain research*, 172(4):533–555.
133. LI, X., LOGAN, R. J. et PASTORE, R. E. (1991). Perception of acoustic source characteristics : Walking sounds. *The Journal of the Acoustical Society of America*, 90(6):3036–3049.
134. LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D. P. et STUDDERT-KENNEDY, M. (1967). Perception of the speech code. *Psychological review*, 74(6):431–461.
135. LIBERMAN, A. M. et MATTINGLY, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21(1):1–36.
136. LINCK, V. (2005). *Modélisation numérique temporelle d'un contact frottant. Mise en évidence d'instabilités locales de contact-Conséquences tribologiques*. Thèse de doctorat, INSA Lyon.
137. LONGCAMP, M., ANTON, J.-L., ROTH, M. et VELAY, J.-L. (2003). Visual presentation of single letters activates a premotor area involved in writing. *NeuroImage*, 19(4):1492–1500.
138. LONGCAMP, M., BOUCARD, C., GILHODES, J.-C., ANTON, J.-L., ROTH, M., NAZARIAN, B. et VELAY, J.-L. (2008). Learning through hand- or typewriting influences visual recognition of new graphic shapes : Behavioral and functional imaging evidence. *Journal of Cognitive Neuroscience*, 20(5):802–815.
139. LONGCAMP, M., TANSKANEN, T. et HARI, R. (2006). The imprint of action : Motor cortex involvement in visual perception of handwritten letters. *NeuroImage*, 33(2):681–688.
140. LONGCAMP, M. et VELAY, J. (2010). *Connaissances motrices, perception visuelle et jugements esthétiques : l'exemple des formes graphiques*, pages 145–154. Champ vallon.
141. LUNEBURG, R. K. (1950). The metric of binocular visual space. *Journal of the Optical Society of America*, 40(10):627–640.
142. LUTFI, R. A. (2007). Human sound source identification. In *Auditory perception of sound sources*, pages 13–42. Springer.
143. MARAVITA, A., SPENCE, C. et DRIVER, J. (2003). Multisensory integration and the body schema : close to hand and within reach. *Current Biology*, 13(13):R531–R539.
144. MCADAMS, S. (1993). Recognition of sound sources and events. In MCADAMS, S. et BIGAND, E., éditeurs : *Thinking in sound : The cognitive psychology of human audition*, Oxford science publications., pages 146–198. New York, NY, US.
145. MCADAMS, S., BEAUCHAMP, J. W. et MENEGUZZI, S. (1999). Discrimination of musical instrument sounds resynthesized with simplified spectrotemporal parameters. *The Journal of the Acoustical Society of America*, 105(2):882–897.
146. MCADAMS, S., CHAIGNE, A. et ROUSSARIE, V. (2004). The psychomechanics of simulated sound sources : Material properties of impacted bars. *The Journal of the Acoustical Society of America*, 115(3):1306–1320.
147. MCADAMS, S. et CUNIBLE, J.-C. (1992). Perception of timbral analogies. *Philosophical transactions : Biological sciences*, pages 383–389.
148. MCADAMS, S., WINSBERG, S., DONNADIEU, S., DE SOETE, G. et KRIMPHOFF, J. (1995). Perceptual scaling of synthesized musical timbres : Common dimensions, specificities, and latent subject classes. *Psychological research*, 58(3):177–192.
149. McDERMOTT, J. H., SCHEMITSCH, M. et SIMONCELLI, E. P. (2013). Summary statistics in auditory perception. *Nature neuroscience*, 16(4):493–498.

150. MCDERMOTT, J. H. et SIMONCELLI, E. P. (2011). Sound texture perception via statistics of the auditory periphery : evidence from sound synthesis. *Neuron*, 71(5):926–940.
151. MERER, A. (2011). *Caractérisation acoustique et perceptive du mouvement évoqué par les sons pour le contrôle de la synthèse*. Thèse de doctorat, Aix-Marseille Université.
152. MERER, A., ARAMAKI, M., YSTAD, S. et KRONLAND-MARTINET, R. (2013). Perceptual characterization of motion evoked by sounds for synthesis control purposes. *ACM Transaction on Applied Perception*, 10(1):1 :1–1 :24.
153. MERER, A., YSTAD, S., KRONLAND-MARTINET, R. et ARAMAKI, M. (2008). Semiotics of sounds evoking motions : Categorization and acoustic features. In KRONLAND-MARTINET, R., YSTAD, S. et JENSEN, K., éditeurs : *Computer Music Modeling and Retrieval. Sense of Sounds*, numéro 4969 de Lecture Notes in Computer Science, pages 139–158. Springer Berlin Heidelberg.
154. MICHAELS, C. F. et CARELLO, C. (1981). *Direct perception*. Prentice-Hall Englewood Cliffs, NJ.
155. MICHEL, C., PISELLA, L., PRABLANC, C., RODE, G. et ROSSETTI, Y. (2007). Enhancing visuomotor adaptation by reducing error signals : single-step (aware) versus multiple-step (unaware) exposure to wedge prisms. *Journal of Cognitive Neuroscience*, 19(2):341–350.
156. MICOULAUD-FRANCHI, J.-A., ARAMAKI, M., MERER, A., CERMOLACCE, M., YSTAD, S., KRONLAND-MARTINET, R., NAUDIN, J. et VION-DURY, J. (2012). Toward an exploration of feeling of strangeness in schizophrenia : Perspectives on acousmatic and everyday listening. *Journal of abnormal psychology*, 121(3):628–640.
157. MICOULAUD-FRANCHI, J.-A., ARAMAKI, M., MERER, A., CERMOLACCE, M., YSTAD, S., KRONLAND-MARTINET, R. et VION-DURY, J. (2011). Categorization and timbre perception of environmental sounds in schizophrenia. *Psychiatry research*, 189(1):149–152.
158. MOORE, B. C. J. et GLASBERG, B. R. (1983). Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *The Journal of the Acoustical Society of America*, 74(3):750–753.
159. MULDER, T. (2007). Motor imagery and action observation : cognitive tools for rehabilitation. *Journal of neural transmission*, 114(10):1265–1278.
160. NELKEN, I. et de CHEVEIGNÉ, A. (2013). An ear for statistics. *Nature neuroscience*, 16(4):381–382.
161. NOË, A. (2004). *Action in perception*. MIT press.
162. NORMAN, J. (2002). Two visual systems and two theories of perception : An attempt to reconcile the constructivist and ecological approaches. *Behavioral and brain sciences*, 25(01):73–96.
163. O'REGAN, J. K. (1992). Solving the "real" mysteries of visual perception : the world as an outside memory. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 46(3):461.
164. O'REGAN, J. K. (2011). *Why red doesn't sound like a bell : Understanding the feel of consciousness*. Oxford University Press.
165. O'REGAN, J. K. et NOË, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and brain sciences*, 24(05):939–973.
166. PALM, G. et AERTSEN, A. (1986). *Brain theory*. Springer-Verlag Berlin.
167. PANTER, P. F. (1965). *Modulation, noise and spectral analysis*. McGraw-Hill.
168. PAPAXANTHIS, C., PAIZIS, C., WHITE, O., POZZO, T. et STUCCHI, N. (2012). The relation between geometry and time in mental actions. *PloS one*, 7(11):e51191.
169. PARISE, C. V., SPENCE, C. et ERNST, M. O. (2012). When correlation implies causation in multisensory integration. *Current Biology*, 22(1):46–49.
170. PARSONS, L. M. (1994). Temporal and kinematic properties of motor behavior reflected in mentally simulated action. *Journal of Experimental Psychology : Human Perception and Performance*, 20(4):709.

171. PEZZULO, G., BARCA, L., BOCCONI, A. L. et BORGHI, A. M. (2010). When affordances climb into your mind : Advantages of motor simulation in a memory task performed by novice and expert rock climbers. *Brain and cognition*, 73(1):68–73.
172. PIAGET, J. et MATALON, B. (1958). Note on the law of the temporal maximum of some optico-geometric illusions. *The American journal of psychology*, 71(1):277–282.
173. PLAMONDON, R. (1998). A kinematic theory of rapid human movements : Part iii. kinetic outcomes. *Biological Cybernetics*, 78(2):133–145.
174. POLLOCK, F. E. et SAPIRO, G. (1997). Constant affine velocity predicts the 1/3 power law of planar motion perception and generation. *Vision Research*, 37(3):347–353.
175. PRINZ, W. (1997). Perception and action planning. *European journal of cognitive psychology*, 9(2):129–154.
176. RAMSEY, R., CUMMING, J. et EDWARDS, M. G. (2008). Exploring a modified conceptualization of imagery direction and golf putting performance. *International Journal of Sport and Exercise Psychology*, 6(2):207–223.
177. REN, Z., YEH, H. et LIN, M. (2010). Synthesizing contact sounds between textured models. In *Virtual Reality Conference (VR), 2010 IEEE*, pages 139–146. IEEE.
178. REPP, B. H. (1987). The sound of two hands clapping : An exploratory study. *The Journal of the Acoustical Society of America*, 81(4):1100–1109.
179. REPP, B. H. et PENEL, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychological research*, 68(4):252–270.
180. REPP, B. H. et SU, Y.-H. (2013). Sensorimotor synchronization : a review of recent research (2006–2012). *Psychonomic bulletin & review*, 20(3):403–452.
181. RICHARDSON, M. J. et FLASH, T. (2002). Comparing smooth arm movements with the two-thirds power law and the related segmented-control hypothesis. *The Journal of Neuroscience*, 22(18):8201–8211.
182. RISSET, J.-C. (1965). Computer study of trumpet tones. *The Journal of the Acoustical Society of America*, 38(5):912–912.
183. RISSET, J.-C. (2010). A propos d’interdisciplinarité-informatique musicale : synthèse, traitement, perception ; musicologie et stic ; oeuvre musicale et mixité. In *Journées d’Informatique Musicale*.
184. RISSET, J.-C. et WESSEL, D. (1999). *Exploration of timbre by analysis and synthesis*, volume 28. Academic Press.
185. RIZZOLATTI, G., FADIGA, L., GALLESE, V. et FOGASSI, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive brain research*, 3(2):131–141.
186. RODGER, M. W. M. et CRAIG, C. (2011). Timing movements to interval durations specified by discrete or continuous sounds. *Experimental brain research*, 214(3):393–402.
187. ROSCH, E. (1973). Natural categories. *Cognitive psychology*, 4(3):328–350.
188. ROSCH, E. (1975). Cognitive representations of semantic categories. *Journal of experimental psychology : General*, 104(3):192–233.
189. ROSCH, E. et MERVIS, C. B. (1975). Family resemblances : Studies in the internal structure of categories. *Cognitive psychology*, 7(4):573–605.
190. ROSSING, T. D. (1994). Acoustics of the glass harmonica. *The Journal of the Acoustical Society of America*, 95(2):1106–1111.
191. ROVAN, J. B., WANDERLEY, M. M., DUBNOV, S. et DEPALLE, P. (1997). Instrumental gestural mapping strategies as expressivity determinants in computer music performance. In *Proceedings of Kansei-The Technology of Emotion Workshop*, pages 3–4.
192. RUNESON, S. (1974). Constant velocity—not perceived as such. *Psychological Research*, 37(1):3–23.
193. SALLAGOÏTY, I., ATHÈNES, S., ZANONE, P.-G. et ALBARET, J.-M. (2004). Stability of coordination patterns in handwriting : Effects of speed and hand. *Motor Control*, 8:405–421.

194. SAVITZKY, A. et GOLAY, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, 36(8):1627–1639.
195. SCHELLENG, J. C. (1973). The bowed string and the player. *The Journal of the Acoustical Society of America*, 53(1):26–41.
196. SCHOMAKER, L. R. B. et PLAMONDON, R. (1990). The relation between pen force and pen-point kinematics in handwriting. *Biological Cybernetics*, 63(4):277–289.
197. SEKULER, R. and Sekuler, A. B. et LAU, R. (1997). Sound alters visual motion perception. *Nature*, 385(6614):308–308.
198. SHAMS, L., KAMITANI, Y. et SHIMOJO, S. (2000). What you see is what you hear. *Nature*, 408(6814):788.
199. SHEPARD, R. N. (1984). Ecological constraints on internal representation : Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological review*, 91(4):417.
200. SHEPARD, R. N. et JUDD, S. A. (1976). Perceptual illusion of rotation of three-dimensional objects. *Science*, 191(4230):952–954.
201. SHEPARD, R. N. et METZLER, J. (1971). Mental rotation of three-dimensional objects.
202. SHOVE, P. et REPP, B. H. (1995). Musical motion and performance : Theoretical and empirical perspectives. *The practice of performance*, pages 55–83.
203. SMITH, M. J. et MERSEREAU, R. M. (1991). *Introduction to digital signal processing : a computer laboratory textbook*.
204. SMITS-ENGELSMAN, B. C. M., NIEMEIJER, A. S. et van GALEN, G. P. (2001). Fine motor deficiencies in children diagnosed as dcd based on poor grapho-motor ability. *Human movement science*, 20(1):161–182.
205. SOKAL, R. R. et ROHLF, F. J. (1962). The comparison of dendrograms by objective methods. *Taxon*, 11(2):33.
206. SOTO-FARACO, S., KINGSTONE, A. et SPENCE, C. (2003). Multisensory contributions to the perception of motion. *Neuropsychologia*, 41(13):1847–1862.
207. SOTO-FARACO, S., LYONS, J., GAZZANIGA, M., SPENCE, C. et KINGSTONE, A. (2002). The ventriloquist in motion : Illusory capture of dynamic information across sensory modalities. *Cognitive brain research*, 14(1):139–146.
208. SOTO-FARACO, S., SPENCE, C. et KINGSTONE, A. (2004). Cross-modal dynamic capture : congruency effects in the perception of motion across sensory modalities. *Journal of Experimental Psychology : Human Perception and Performance*, 30(2):330.
209. SPENCE, C. (2011). Crossmodal correspondences : A tutorial review. *Attention, Perception, & Psychophysics*, 73(4):971–995.
210. STOIMENOV, B., MARUYAMA, S., ADACHI, K. et KATO, K. (2007). The roughness effect on the frequency of frictional sound. *Tribology International*, 40(4):659–664.
211. STUCCHI, N. et VIVIANI, P. (1993). Cerebral dominance and asynchrony between bimanual two-dimensional movements. *Journal of Experimental Psychology : Human Perception and Performance*, 19(6):1200.
212. TERRIEN, S., VERGEZ, C. et FABRE, B. (2013). Flute-like musical instruments : a toy model investigated through numerical continuation. *Journal of sound and vibration*, 332(15):3833–3848.
213. TEULINGS, H.-L., THOMASSEN, A. J. W. M. et van GALEN, G. P. (1986). Invariants in handwriting : The information contained in a motor program. *Graphonomics : Contemporary research in handwriting*, pages 305–315.
214. THEUREL, A., FRILEUX, S., HATWELL, Y. et GENTAZ, E. (2012). The haptic recognition of geometrical shapes in congenitally blind and blindfolded adolescents : is there a haptic prototype effect ? *PloS one*, 7(6):e40251.
215. THORET, E., ARAMAKI, M., BOURDIN, C., BRINGOUX, L., KRONLAND-MARTINET, R. et YSTAD, S. Audio-motor synchronization : the effect of mapping between kinematics and acoustic cues on geometric motor features. In *Sound, Music and Motion, LNCS 8905*, pages 234–245.
216. THORET, E., ARAMAKI, M., BRINGOUX, L., KRONLAND-MARTINET, R. et YSTAD, S. (2014a). When acoustic stimuli turn visual circles into ellipses : sounds evoking accelerations modify visuo-motor coupling. *15th International Multisensory Research Forum (IMRF), Amsterdam, Jun. 11-14, 2014*.

217. THORET, E., ARAMAKI, M., GONDRE, C., KRONLAND-MARTINET, R. et YSTAD, S. Controlling a non linear friction model for evocative sound synthesis applications. In *International Conference on Digital Audio Effects (DAFx)*, Maynooth, Ireland, Sept. 2013.
218. THORET, E., ARAMAKI, M., KRONLAND-MARTINET, R., VELAY, J.-L. et YSTAD, S. (2012a). From shape to sound : sonification of two dimensional curves by reenaction of biological movements. In *Proceedings of the 9th International Symposium on Computer Music Modeling and Retrieval (CMMR)*, 19-22 June, London, 2012.
219. THORET, E., ARAMAKI, M., KRONLAND-MARTINET, R., VELAY, J.-L. et YSTAD, S. (2012b). Sonifying drawings : characterization of perceptual attributes of sounds produced by human gestures. In *Acoustics 2012, Nantes*, pages 1089–1094.
220. THORET, E., ARAMAKI, M., KRONLAND-MARTINET, R., VELAY, J.-L. et YSTAD, S. (2013). Reenacting sensorimotor features of drawing movements from friction sounds. In *From Sounds to Music and Emotions, LNCS 7900*, pages 130–153. Springer.
221. THORET, E., ARAMAKI, M., KRONLAND-MARTINET, R., VELAY, J.-L. et YSTAD, S. (2014b). Auditory perception of biological movements : an evidence of cognitive specificities from sound synthesis. In *15th International Multisensory Research Forum (IMRF)*, Amsterdam, Jun. 11-14, 2014.
222. THORET, E., ARAMAKI, M., KRONLAND-MARTINET, R., VELAY, J.-L. et YSTAD, S. (2014c). From sound to shape : Auditory perception of drawing movements. *Journal of Experimental Psychology : Human Perception and Performance*, 40(3):983–994.
223. THORNTON, I. M. et KNOBLICH, G. (2006). Action perception : seeing the world through a moving body. *Current Biology*, 16(1):R27–R29.
224. van den DOEL, K. (2005). Physically based models for liquid sounds. *ACM Transactions on Applied Perception (TAP)*, 2(4):534–546.
225. VAN DEN DOEL, K., KRY, P. G. et PAI, D. K. (2001). FoleyAutomatic : physically-based sound effects for interactive simulation and animation. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 537–544.
226. VARELA, F. J. (1996). *Invitation aux sciences cognitives*. Seuil.
227. VARELA, F. J., THOMPSON, E. et ROSCH, E. (1993). *L'inscription corporelle de l'esprit*. Seuil.
228. VERRON, C., ARAMAKI, M., KRONLAND-MARTINET, R. et PALLONE, G. (2010). A 3-d immersive synthesizer for environmental sounds. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(6):1550–1561.
229. VIVIANI, P. (1986). Do units of motor action really exist? *Experimental Brain Research*, 15:201–215.
230. VIVIANI, P., BAUD-BOVY, G. et REDOLFI, M. (1997). Perceiving and tracking kinesthetic stimuli : Further evidence of motor-perceptual interactions. *Journal of Experimental Psychology : Human Perception and Performance*, 23(4):1232.
231. VIVIANI, P., CAMPADELLI, P. et MOUNOUD, P. (1987). Visuo-manual pursuit tracking of human two-dimensional movements. *Journal of Experimental Psychology : Human Perception and Performance*, 13(1):62.
232. VIVIANI, P. et FLASH, T. (1995). Minimum-jerk, two-thirds power law, and isochrony : converging approaches to movement planning. *Journal of American Psychology*, 21:32–53.
233. VIVIANI, P. et MCCOLLUM, G. (1983). The relation between linear extent and velocity in drawing movements. *Neuroscience*, 10(1):211–218.
234. VIVIANI, P. et MOUNOUD, P. (1990). Perceptuomotor compatibility in pursuit tracking of two-dimensional movements. *Journal of motor behavior*, 22(3):407–443.
235. VIVIANI, P. et SCHNEIDER, R. (1991). A developmental study of the relationship between geometry and kinematics in drawing movements. *Journal of Experimental Psychology*, 17(1):198–218.
236. VIVIANI, P. et STUCCHI, N. (1989). The effect of movement velocity on form perception : Geometric illusions in dynamic displays. *Perception and Psychophysics*, 46(3):266–274.

237. VIVIANI, P. et STUCCHI, N. (1992). Biological movements look uniform : Evidence of motor-perceptual interactions. *Journal of Experimental Psychology : Human Perception and Performance*, 18(3):603–623.
238. VIVIANI, P. et TERZUOLO, C. (1982). Trajectory determines movement dynamics. *Neuroscience*, 7(2):431–437.
239. VON HELMHOLTZ, H. (1912). *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. Longmans, Green.
240. WAMAIN, Y., TALLET, J., ZANONE, P.-G. et LONGCAMP, M. (2011). “biological geometry perception” : Visual discrimination of eccentricity is related to individual motor preferences. *PloS one*, 6(1):e15995.
241. WARREN, W. H. et VERBRUGGE, R. R. (1984). Auditory perception of breaking and bouncing events : A case study in ecological acoustics. *Journal of Experimental Psychology : Human Perception and Performance*, 10(5):704–712.
242. WEISS, Y., SIMONCELLI, E. P. et ADELSON, E. H. (2002). Motion illusions as optimal percepts. *Nature neuroscience*, 5(6):598–604.
243. WILDES, R. P. et RICHARDS, W. A. (1988). Recovering material properties from sound. *Natural computation*, pages 356–363.
244. YOUNG, D. et SERAFIN, S. (2003). Playability evaluation of a virtual bowed string instrument. In *Proceedings of the 2003 conference on New Interfaces for Musical Expression (NIME)*, pages 104–108. National University of Singapore.
245. YOUNG, W., RODGER, M. et CRAIG, C. M. (2013). Perceiving and reenacting spatiotemporal characteristics of walking sounds. *Journal of Experimental Psychology : Human Perception and Performance*, 39(2):464–476.
246. ZANONE, P. G. et KELSO, J. A. (1992). Evolution of behavioral attractors with learning : nonequilibrium phase transitions. *Journal of Experimental Psychology : Human Perception and Performance*, 18(2):403.
247. ZATORRE, R. J. et HALPERN, A. R. (2005). Mental concerts : musical imagery and auditory cortex. *Neuron*, 47(1):9–12.
248. ZMIGROD, S. et HOMMEL, B. (2013). Feature integration across multimodal perception and action : a review. *Multisensory Research*, 26:143–157.

Titre : Caractérisation acoustique des relations entre les mouvements biologiques et la perception sonore : application au contrôle de la synthèse et à l'apprentissage de gestes

Résumé : Cette thèse s'est intéressée aux relations entre les mouvements biologiques et la perception sonore en considérant le cas spécifique des mouvements graphiques et des sons de frottement qu'ils génèrent. L'originalité de ces travaux réside dans l'utilisation d'un modèle de synthèse sonore basé sur un principe perceptif issu de l'approche écologique de la perception et contrôlé par des modèles de gestes. Des stimuli sonores dont le timbre n'est modulé que par des variations de vitesse produites par un geste ont ainsi pu être générés permettant de se focaliser sur l'influence perceptive de cet invariant transformationnel. Une première étude a ainsi montré que l'on reconnaît la cinématique des mouvements biologiques (la loi en puissance $1/3$), et que l'on peut discriminer des formes géométriques simples juste à partir des sons de frottement produits. Une seconde étude a montré l'existence de prototypes dynamiques sonores caractérisant les trajectoires elliptiques, mettant ainsi en évidence que les prototypes géométriques peuvent émerger d'un couplage sensorimoteur. Enfin, une dernière étude a montré qu'une cinématique évoquée par un sonore influence significativement la cinématique et la géométrie d'un geste dans une tâche de reproduction graphique du mouvement d'un point lumineux. Ce résultat révèle l'importance de la modalité auditive dans l'intégration multisensorielle des mouvements continus dans une situation jamais explorée. Ces résultats ont permis le contrôle de modèles de synthèse par des descriptions gestuelles et la création d'outils de sonification pour l'apprentissage de gestes et la réhabilitation d'une pathologie motrice, la dysgraphie.

Mots-clés : Perception sonore – Relation geste/son – Mouvements biologiques – Synthèse sonore – Contrôle intuitif – Analyse par la synthèse – Synchronisation sensori-motrice – Perception multisensorielle du mouvement – Sonification

Title : Acoustic characterisation of relations between biological movements and auditory perception : applications to the control of sound synthesis and gesture learning

Summary : This thesis focused on the relations between biological movements and auditory perception in considering the specific case of graphical movements and the friction sounds they produced. The originality of this work lies in the use of sound synthesis processes that are based on a perceptual paradigm and that can be controlled by gesture models. The present synthesis model made it possible to generate acoustic stimuli which timbre was directly modulated by the velocity variations induced by a graphic gesture in order to exclusively focus on the perceptual influence of this transformational invariant. A first study showed that we can recognize the biological motion kinematics (the $1/3$ power law) and discriminate simple geometric shapes simply by listening to the timbre variations of friction sounds that solely evoke velocity variations. A second study revealed the existence of dynamic prototypes characterized by sounds corresponding to the most representative elliptic trajectory, thus revealing that prototypical shapes may emerged from sensorimotor coupling. A final study showed that the kinematics evoked by friction sounds may significantly affect the dynamic and geometric dimension in the visuo-motor coupling. This shed critical light on the relevance of auditory perception in the multisensory integration of continuous motion in a situation never explored. All of these theoretical results enabled the gestural control of sound synthesis models from a gestural description and the creation of sonification tools for gesture learning and rehabilitation of a graphomotor disease, dysgraphia.

Keywords : Auditory perception – Sound/gesture relation – Biological motions – Sound synthesis – Analysis-synthesis – Sensorimotor synchronization – Multisensory perception of motion – Sonification